

# Diffusion Approximations for Double-ended Queues with Reneging in Heavy Traffic

Xin Liu

July 18, 2016

**Abstract:** We study a double-ended queue which consists of two classes of customers. Whenever there is a pair of customers from both classes, they are matched and leave the system immediately. The matching follows first-come-first-serve principle. If a customer from one class cannot be matched immediately, he/she will stay in a queue and wait for the upcoming arrivals from the other class. Thus there cannot be non-zero numbers of customers from both classes simultaneously in the system. We also assume that each customer can leave the queue without being matched because of impatience. The arrival processes are assumed to be independent renewal processes, and the patience times for both classes are generally distributed. Under suitable heavy traffic conditions, assuming that the diffusion-scaled queue length process is stochastically bounded, we establish a simple asymptotic relationship between the diffusion-scaled queue length process and the diffusion-scaled offered waiting time processes, and further show that the diffusion-scaled queue length process converges weakly to a diffusion process. We also provide a sufficient condition for the stochastic boundedness of the diffusion-scaled queue length process. At last, the explicit form of the stationary distribution of the limit diffusion process is derived.

**AMS 2010 subject classifications:** Primary: 60F05, 60K25, 90B22; Secondary: 60K05.

**Keywords:** Double-ended queues; Matching systems; First-Come-First-Serve; Customer abandonment; Generally distributed patience times; Hazard rate scaling; Diffusion approximations; Stationary distribution; Heavy traffic.

## 1. Introduction

Consider a simple matching system which consists of two classes of customers. Whenever there is a pair of customers from both classes, they are matched and leave the system immediately. The matching follows first-come-first-serve principle. If a customer from one class cannot be matched immediately, he/she will stay in a queue and wait for the upcoming arrivals from the other class. Thus there cannot be non-zero numbers of customers from both classes simultaneously in the system. Customers are assumed to be impatient and they can leave the system without being matched. Such system forms a double-ended queueing system, which is schematically shown in Figure 1. We assume that the arrival processes are independent renewal processes, and the patience times for customers of each class are IID with a general distribution. Under certain conditions, we establish a simple asymptotic relationship between the diffusion-scaled queue length process and the diffusion-scaled offered waiting time processes, and show that the diffusion-scaled queue length process converges weakly to a diffusion process. Those conditions consist of a suitable heavy traffic condition, a mild condition on patience-time distributions, and an assumption that the diffusion-scaled queue length process is stochastically

bounded. We further show that under the heavy traffic condition and a proper scaling of the hazard rate functions of the patience-time distributions, the diffusion-scaled queue length is stochastically bounded. Finally, the unique stationary distribution of the diffusion limit is derived. Diffusion approximations for double-ended queues with renewal arrivals and exponential patience times are studied in [20], and our current work extends the heavy traffic diffusion approximation result in [20] to a setting with generally distributed patience times.

Double-ended queues have been studied for many applications, including taxi-service systems, financial markets, assembly systems, perishable inventory systems, and organ transplant systems. Double-ended queues with Poisson arrivals were first introduced by Kendall in [17], and Dobbie [9] studied the probability of a given queue size of double-ended queues proposed by Kendall. Later on, Giveen [12] considered double-ended queues with time-dependent Poisson arrivals, and studied the limiting distributions, and Kashyap [16] studied a taxi service system with limited waiting space, where the arrivals of taxis and passengers are assumed to be Poisson processes, and derived the steady state distribution. In [3], Conolly et al. studied the impatience behavior of customers under the assumption of exponential arrivals, services, and patience times. Applications of double-ended queues on financial markets, networks with synchronization nodes, perishable inventory systems, and organ transplant systems can be found in [1, 30, 24, 23, 2], etc. When renewal arrivals and/or generally distributed patience times are considered, the explicit form of the limiting distribution becomes intractable. Degirmenci [8] studied the asymptotic behaviors of the Markovian double-ended queue, and conjectured that the result holds for the setting with renewal arrivals and general patience times. In [18], Kim et al. studied the extended double-ended queue which allowed bulk arrivals, general patience times, positive processing times, etc, using simulation methods. Zenios [30] applied Laplace's method to a double-ended queue where the inter-arrivals of supply were generally distributed, and demand arrived according a Poisson process, and only demand had abandonments with generally distributed patience times. In this work, we develop rigorous diffusion approximations for double-ended queues with renewal arrivals and generally distributed patience times under appropriate asymptotic regimes. Diffusion approximations of (one-sided) queueing systems with reneging have been studied in both conventional heavy traffic regime and Halfin-Whitt regime. Roughly speaking, in both regimes, the queueing system is nearly critically loaded. The conventional heavy traffic analysis can be found in [27, 28, 26, 19], while [11, 29, 5, 21, 4, 25] study many-server queues in Halfin-Whitt regime.

In this work, our main assumption on patience time distributions is similar to that considered by Lee and Weerasinghe [19], which requires the patience time distribution functions satisfy some scaling limit results, and are more general than the commonly used assumptions in the above literatures (see Assumption 3.3 and Remark 3.1 (iii)). Our main results, Theorems 4.1 and 4.2, are established under this assumption, the heavy traffic assumption (Assumption 3.2), and the assumption that the diffusion-scaled queue length process is stochastically bounded. We then consider the hazard rate scaling as in [26, 25, 15] (see Assumption 3.4), and show in Corollary 4.4 that under such hazard rate scaling and the heavy traffic condition, the diffusion-scaled queue length process is stochastically bounded and so all the results hold. It is worth noting that the condition that the diffusion-scaled queue length process is stochastically bounded is considered by Dai and He [4], and they establish the diffusion approximations for a  $G/G/n + GI$  queue under this condition, and a more strict condition on patience time distribution in Halfin-Whitt regime. Furthermore, Reed and Tezcan [25] obtain a heavy traffic limit for  $GI/M/n + GI$  queue by assuming that the patience time distribution satisfies an

appropriate hazard rate scaling.

The rest of the paper is organized as follows. In Section 2, we introduce the model of the double-ended queue with renewal arrivals and generally distributed patience times. In Section 3, we present the asymptotic framework by introducing the main assumptions. All the main results are collected in Section 4, and their proofs are provided in Section 5, and Appendix.

We use the following notation. Denote by  $\mathbb{R}$ ,  $\mathbb{R}_+$ ,  $\mathbb{Z}$ , and  $\mathbb{N}$  the sets of real numbers, non-negative real numbers, integers, and positive integers, respectively. For a real number  $a$ , define  $a^+ = \max\{a, 0\}$  and  $a^- = \max\{0, -a\}$ . Similarly, for a real function  $f$  defined on  $[0, \infty)$ , define  $f^+(t) = \max\{0, f(t)\}$  and  $f^-(t) = \max\{0, -f(t)\}$ ,  $t \geq 0$ . Denote by  $\mathcal{D}([0, \infty); \mathbb{R})$  the space of right continuous functions with left limits defined from  $[0, \infty)$  to  $\mathbb{R}$  with the usual Skorohod topology. For  $x \in \mathcal{D}([0, \infty); \mathbb{R})$ , let

$$\|x\|_t = \sup_{s \in [0, t]} |x(s)|, \quad t \geq 0,$$

and

$$\|x\|_{s,t} = \sup_{u \in [s, t]} |x(u)|, \quad t \geq s \geq 0.$$

A function  $f : [0, \infty) \rightarrow [0, \infty)$  is called locally Lipschitz continuous if for any  $t \in [0, \infty)$ , there exists  $\kappa \in (0, \infty)$  (may depend on  $t$ ) such that for  $x_1, x_2 \in [0, t]$ ,

$$|f(x_1) - f(x_2)| \leq \kappa |x_1 - x_2|.$$

For a stochastic process  $X$ , we will use the notation  $X(t)$  and  $X_t$  interchangeably. For a semimartingale  $Y \in \mathcal{D}([0, \infty); \mathbb{R})$ , we denote by  $[Y]$  the quadratic variation of  $Y$ . A sequence of stochastic processes  $\{X^n\}$  is called stochastically bounded if for each  $T \geq 0$ ,

$$\lim_{K \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P}(\|X^n\|_T > K) = 0.$$

## 2. Network model

We study a double-ended queue which consists of two classes of customers – Classes 1 and  $-1$ . Whenever there is a pair of customers from both classes, they are matched and leave the system immediately. The matching follows first-come-first-served principle. If a customer from one class cannot be matched immediately, he/she will stay in a queue and wait for the upcoming arrivals from the other class. We also assume that each customer is impatient. A double-ended queue is schematically given in Figure 1.

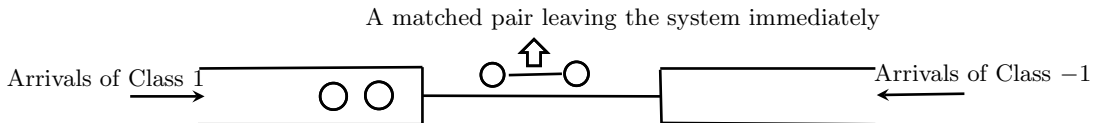


FIG 1. A double-ended Queue.

Let  $Q(t)$  be the length of the double-ended queue at time  $t$ . Different from queue length processes in one-sided queueing systems, here  $Q(t) \in \mathbb{Z}$ , and we let  $Q^+(t)$  denote the number of customers of Class 1 waiting in the queue at time  $t$ , and  $Q^-(t)$  the number of customers of Class  $-1$  waiting in the queue at time  $t$ . For  $t \geq 0$ , let

$$Q_1(t) = Q^+(t) \quad \text{and} \quad Q_{-1}(t) = Q^-(t).$$

Without loss of generality, we assume  $Q_1(0) \geq 0$  and  $Q_{-1}(0) = 0$ .

Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a complete probability space satisfying the usual conditions. All the random variables and stochastic processes are assumed to be defined on this space. Let  $i = \pm 1$ . We assume the inter-arrival times for customers of Class  $i$  are i.i.d. random variables  $\{u_{i,k} : k \in \mathbb{N}\}$ , and  $u_{i,1}$  has mean  $\frac{1}{\lambda_i}$  and standard deviation  $\sigma_i$ . Define

$$N_i(t) = \max \left\{ k \in \mathbb{N} : \sum_{l=1}^k u_{i,l} \leq t \right\}.$$

The renewal process  $N_i$  is the arrival process for Class  $i$ . We assume that the patience times of customers of Class  $i$  are given by an i.i.d. sequence  $\{d_{i,k} : k \in \mathbb{Z}\}$ . For  $k \in \mathbb{N}$ ,  $d_{i,k}$  represents the patience time of the  $k$ -th customer of Class  $i$  who enters the system after time 0, and for  $k \in -\mathbb{N} \cup \{0\}$ ,  $d_{i,k}$  is the patience time of the  $(-k+1)$ -st customer of Class  $i$  who enters the system prior to time 0 (if such customers exist). We assume that  $d_{i,1}$  has cumulative distribution function  $F_i$ . Finally, we assume  $Q(0)$ ,  $\{u_{1,k} : k \in \mathbb{N}\}$ ,  $\{u_{-1,k} : k \in \mathbb{N}\}$ ,  $\{d_{1,k} : k \in \mathbb{Z}\}$ , and  $\{d_{-1,k} : k \in \mathbb{Z}\}$  are independent.

For  $k \in \mathbb{Z}$ , define  $t_{i,k}$  to be the arrival time of the  $k^{\text{th}}$  customer of Class  $i$ . More precisely, for  $k \in \mathbb{N}$ ,  $t_{i,k} = \sum_{l=1}^k u_{i,l}$ , and for  $k \in -\mathbb{N} \cup \{0\}$ , we let  $t_{i,k} = 0$ , that is all customers, who arrive prior to time 0, have arrival times equal to 0. For  $k \in \mathbb{Z}$ , let  $w_{i,k}$ , which is called the *offered waiting time*, denote the waiting time that the  $(k1_{\{k>0\}} + (-k+1)1_{\{k \leq 0\}})$ -th customer of Class  $i$  needs to experience assuming her/his patience time is infinite. For a customer who arrives prior to time 0, we assume her/his waiting time starts to count at time 0. Let  $G_i(t)$  denote the number of Class  $i$  customers who abandon the system by time  $t$ . Then for  $t \geq 0$ ,

$$G_i(t) = \sum_{k=-Q_i(0)+1}^{N_i(t)} 1_{\{d_{i,k} \leq w_{i,k}, t_{i,k} + d_{i,k} \leq t\}}. \quad (2.1)$$

The process  $G_i$  will be called the *abandonment process* for Class  $i$ . The *queue length process* now can be formulated as follows:

$$Q(t) = Q(0) + N_1(t) - N_{-1}(t) - G_1(t) + G_{-1}(t). \quad (2.2)$$

To study the offered waiting times, we next define  $R_i(t)$  to be the number of Class  $i$  customers who have arrived by time  $t$ , and will abandon the system eventually. Then for  $t \geq 0$ ,

$$R_i(t) = \sum_{k=-Q_i(0)+1}^{N_i(t)} 1_{\{d_{i,k} \leq w_{i,k}\}}. \quad (2.3)$$

For Class 1, there are  $Q_1(0) = Q^+(0)$  customers entering the system prior to time 0, and for such customers, define

$$R_{1,k}(0) = \sum_{l=k+1}^0 1_{\{d_{1,l} \leq w_{1,l}\}}, \quad k = -Q_1(0) + 1, \dots, -1, \text{ and } R_{1,0} = 0.$$

Here  $R_{1,k}, k = -Q_1(0) + 1, \dots, -1, 0$ , represents the number of the first  $-k$  customers of Class 1 who arrives prior to time 0, and will abandon the system eventually. In the following, we characterize the offered waiting times of customers in terms of  $\{t_{-i,k}, k \in \mathbb{Z}\}$ ,  $\{R_i(t), t \geq 0\}$  and  $\{R_{1,k}(0), k = -Q_1(0) + 1, \dots, -1, 0\}$ . First, it is clear that the offered waiting time of the first customer of Class 1 who arrives prior to time 0 is  $t_{-1,1}$  (i.e., the arrival time of the first customer of Class  $-1$ ). For  $k = -Q_1(0) + 1, \dots, -1$ , assuming he/she is patient, the  $(-k + 1)$ -st customer of Class 1 arriving before time 0 will be matched with the  $(-k + 1 - R_{1,k}(0))$ -th customer of Class  $-1$ . Thus the offered waiting time of the  $(-k + 1)$ -st customer of Class 1 who arrives before time 0 is

$$w_{1,k} = t_{-1, -k+1-R_{1,k}(0)}. \quad (2.4)$$

We next consider the customers arriving after time 0. For  $k \in \mathbb{N}$ , at time  $t_{1,k}$ , the  $k$ -th customer of Class 1 arrives at the system. If  $Q_{-1}(t_{1,k}-) > 0$  (i.e., there are customers of Class  $-1$  waiting at time  $t_{1,k}$ ), then the offered waiting time is

$$w_{1,k} = 0. \quad (2.5)$$

If  $Q_{-1}(t_{1,k}-) = 0$  (i.e., there is no customer of Class  $-1$  waiting at time  $t_{1,k}$ ), this  $k$ -th customer of Class 1 must wait in the system for the  $(k + Q_1(0) - R_1(t_{1,k}-) + G_{-1}(t_{1,k}-))$ -th customer of Class  $-1$ , and the offered waiting time is

$$w_{1,k} = t_{-1, k+Q_1(0)-R_1(t_{1,k}-)+G_{-1}(t_{1,k}-)} - t_{1,k}.$$

Next noting that  $R_{-1}(t) = G_{-1}(t)$  when  $Q_{-1}(t) = 0$ , we have

$$w_{1,k} = t_{-1, k+Q_1(0)-R_1(t_{1,k}-)+R_{-1}(t_{1,k}-)} - t_{1,k}. \quad (2.6)$$

Combining (2.4), (2.5) and (2.6), we have that

$$w_{1,k} = \begin{cases} \left( t_{-1, k+Q_1(0)-R_1(t_{1,k}-)+R_{-1}(t_{1,k}-)} - t_{1,k} \right) 1_{\{Q_{-1}(t_{1,k}-)=0\}}, & k \in \mathbb{N}, \\ t_{-1, -k+1-R_{1,k}(0)}, & -Q_1(0) + 1 \leq k \leq 0. \end{cases} \quad (2.7)$$

Using similar arguments, we have for Class  $-1$ ,

$$w_{-1,k} = \left( t_{1, k-R_{-1}(t_{-1,k}-)-Q_1(0)+R_1(t_{-1,k}-)} - t_{-1,k} \right) 1_{\{Q_1(t_{-1,k}-)=0\}}, \quad k \in \mathbb{N}. \quad (2.8)$$

It is shown in Lemma 5.1 that for  $i = \pm 1$  and  $k \in \mathbb{N}$ ,

$$w_{i,k} = \left[ t_{-i, k+Q_i(0)-R_i(t_{i,k}-)-Q_{-i}(0)+R_{-i}(t_{i,k}-)} - t_{i,k} \right]^+. \quad (2.9)$$

We thus define the virtual waiting time process  $W_i(t)$ , which can be interpreted as the amount of time that a hypothetical customer of Class  $i$  arriving right after time  $t$  will have to wait in the system, as follows. For  $t \geq 0$ ,

$$W_i(t) = \left[ t_{-i, N_i(t)+1+Q_i(0)-R_i(t)-Q_{-i}(0)+R_{-i}(t)} - t \right]^+. \quad (2.10)$$

It is clear that  $W_i(t_{i,k}-) = w_{i,k}$  for  $k \in \mathbb{N}$ .

**Remark 2.1.**

- (i) In (2.8), it is possible that the subscript  $k - R_{-1}(t_{-1,k-}) - Q_1(0) + R_1(t_{-1,k-}) \leq 0$ . When this happens, we have  $Q_1(t_{-1,k-}) > 0$ , and  $w_{-1,k} = 0$ . (See Lemma 5.1 for its proof.) In fact, this  $k$ -th customer of Class  $-1$  will be matched with a customer of Class 1 who arrives prior to time 0 and are waiting in the queue at time  $t_{-1,k}$ .
- (ii) Lemma 5.1 shows that for  $k \in \mathbb{N}$ , in (2.7), when  $Q_{-1}(t_{1,k-}) = 0$ , we have

$$t_{-1,k+Q_1(0)-R_1(t_{1,k-})+R_{-1}(t_{1,k-})} \geq t_{1,k},$$

and in (2.8), when  $k - R_{-1}(t_{-1,k-}) - Q_1(0) + R_1(t_{-1,k-}) > 0$ , and  $Q_1(t_{-1,k-}) = 0$ , we have

$$t_{1,k-R_{-1}(t_{-1,k-})-Q_1(0)+R_1(t_{-1,k-})} \geq t_{-1,k}.$$

**3. Asymptotic framework**

We consider a sequence of double-ended queues indexed by  $n \in \mathbb{N}$ . For the  $n$ -th system, all the notation introduced in Section 2 is carried forward except that we append a superscript  $n$  to all quantities to indicate the dependence of parameters, random variables, and stochastic processes on  $n$ . Let  $i = \pm 1$ . In particular, we assume all the random variables  $\{u_{i,k}^n : k \in \mathbb{N}\}$ ,  $\{d_{i,k}^n : k \in \mathbb{Z}\}$ ,  $\{v_{i,k}^n : k \in \mathbb{Z}\}$ ,  $\{w_{i,k}^n : k \in \mathbb{Z}\}$ , and stochastic processes  $Q^n, N_i^n, G_i^n, R_i^n, W_i^n$  are defined on the space  $(\Omega^n, \mathcal{F}^n, \mathbb{P}^n)$ . The expectation operator with respect to  $\mathbb{P}^n$  will be denoted by  $\mathbb{E}^n$ . Finally, the cumulative distribution function of  $d_{i,1}^n$  is  $F_i^n$ , and the mean and standard deviation of  $u_{i,1}^n$  are  $\frac{1}{\lambda_i^n}$  and  $\sigma_i^n$ .

The following are our main assumptions.

**Assumption 3.1** (Conditions on inter-arrival times). *There exist independent sequences of i.i.d. nonnegative random variables  $\{\tilde{u}_{i,k}^n : k \in \mathbb{N}\}$ ,  $i = \pm 1$ , such that for  $k \in \mathbb{N}$  and  $n \in \mathbb{N}$ ,*

- (i)  $u_{i,k}^n = \frac{\tilde{u}_{i,k}^n}{n}$ ;
- (ii)  $\mathbb{E}(\tilde{u}_{i,k}^n) \rightarrow \frac{1}{\lambda_i}$ ,  $\text{Var}(\tilde{u}_{i,k}^n) \rightarrow \sigma_i^2$  as  $n \rightarrow \infty$ ;
- (iii)  $\{(\tilde{u}_{i,1}^n)^2 : n \in \mathbb{N}\}$  is uniformly integrable.

**Assumption 3.2** (Heavy traffic condition). *There exists  $c \in \mathbb{R}$  such that*

$$\frac{\lambda_1^n - \lambda_{-1}^n}{\sqrt{n}} \rightarrow c, \text{ as } n \rightarrow \infty. \quad (3.1)$$

**Assumption 3.3** (Conditions on patience times). *For  $i = \pm 1$ ,  $F_i^n(0) = 0$ , and there exists nonnegative nondecreasing locally Lipschitz continuous functions  $H_i$  on  $[0, \infty)$  such that for  $t \geq 0$ ,*

$$\sqrt{n}F_i^n\left(\frac{t}{\sqrt{n}}\right) \rightarrow H_i(t), \text{ as } n \rightarrow \infty.$$

*In particular,  $F_i^n(\frac{t}{\sqrt{n}}) \rightarrow 0$ , as  $n \rightarrow \infty$ .*

**Assumption 3.4** (Conditions on hazard rate functions). *For  $i = \pm 1$ ,  $F_i^n$  has density function  $f_i^n$  and hazard rate function  $h_i^n = f_i^n/(1-F_i^n)$  on  $[0, M_i^n)$ , where  $M_i^n = \sup\{s \geq 0 : F_i^n(s) < 1\}$ . Furthermore, for  $t \in [0, M_i^n)$ ,  $\sup_{n \in \mathbb{N}} \|h_i^n\|_t < \infty$ , and  $h_i^n(t/\sqrt{n}) \rightarrow h_i(t)$  as  $n \rightarrow \infty$ , where  $h_i$  is a nonnegative measurable function.*

**Remark 3.1.**

- (i) From Assumption 3.1, we have for  $i = \pm 1$ ,

$$\lambda_i^n = \frac{1}{\mathbb{E}(u_{i,1}^n)}, \text{ and } \frac{\lambda_i^n}{n} \rightarrow \lambda_i, \quad (3.2)$$

and

$$(\sigma_i^n)^2 = \text{Var}(u_{i,1}^n), \text{ and } n^2(\sigma_i^n)^2 \rightarrow \sigma_i^2. \quad (3.3)$$

- (ii) The heavy traffic condition in Assumption 3.2 implies that  $\lambda_1 = \lambda_{-1}$ . We are going to let

$$\lambda = \lambda_1 = \lambda_{-1}. \quad (3.4)$$

- (iii) Assumption 3.3 is similar to Assumption 3.4 in [19]. We list some examples of the cumulative distribution function satisfying Assumption 3.3 (also see Remark 3.4 of [19]).

- (a) Let  $F$  be a cumulative distribution function which is right differentiable at 0, and  $F(0) = 0$ . Then  $F^n \equiv F$  satisfies Assumption 3.3 with  $H(x) = F'(0)x$ . Such patience times distribution has been considered in [29, 21, 5, 4].
- (b) Consider  $F^n(x) = 1 - \exp\{-\int_0^x h(\sqrt{n}u)du\}$ , where  $h$  is a nonnegative continuous function. Then  $F^n$  satisfies Assumption 3.3 with  $H(x) = \int_0^x h(u)du$ . Here  $h^n(x) \equiv h(\sqrt{n}x)$ ,  $x \geq 0$ , is the hazard rate function of  $F^n$ , and such hazard rate scaling of patience times distribution has been studied in [26, 25, 15].
- (c) There are many other distribution functions besides those in (a) and (b) satisfying Assumption 3.3. For example, let

$$F_i^n(x) = \begin{cases} F^n(x), & x \in [0, \delta) \\ 1, & x \in [\delta, \infty), \end{cases}$$

where  $F^n$  is as in (a) or (b), and  $\delta > 0$ . It is easy to see that  $F_i^n$  has the same scaling limit function as  $F^n$ . In fact, Assumption 3.3 only requires  $F_i^n$  has the scaling limit over a small interval  $[0, \mathcal{O}(\frac{1}{\sqrt{n}})]$ .

- (iv) Under Assumption 3.4, we have  $F_i^n(t) = 1 - \exp\{-\int_0^t h_i^n(u)du\}$ , and  $F_i^n$  satisfies Assumption 3.3 with  $H_i(x) = \int_0^x h_i(u)du$ . Thus the patience times distribution discussed in (iii)(b) is a special case of Assumption 3.4. This assumption is the same as the main assumption in [15].

**4. Main results**

Define the fluid and diffusion-scaled processes as follows. For  $i = \pm 1$  and  $t \geq 0$ , *Fluid scaled processes*:

$$\bar{Q}^n(t) = \frac{Q^n(t)}{n}, \quad \bar{N}_i^n(t) = \frac{N_i^n(t)}{n}, \quad \bar{G}_i^n(t) = \frac{G_i^n(t)}{n}, \quad \bar{R}_i^n(t) = \frac{R_i^n(t)}{n},$$

Diffusion scaled processes:

$$\begin{aligned}\hat{Q}^n(t) &= \frac{Q^n(t)}{\sqrt{n}}, \quad \hat{N}_i^n(t) = \frac{N_i^n(t) - \lambda_i^n t}{\sqrt{n}}, \quad \hat{G}_i^n(t) = \frac{G_i^n(t)}{\sqrt{n}}, \quad \hat{R}_i^n(t) = \frac{R_i^n(t)}{\sqrt{n}}, \\ \hat{W}_i^n(t) &= \sqrt{n} W_i^n(t).\end{aligned}$$

Our main results are presented as follows. Theorem 4.1 establishes an asymptotic linear relationship between the diffusion-scaled offered waiting time processes and the diffusion scaled queue length process for both classes. The scaling limit theorem for the queue length process is provided in Theorem 4.2, and we derive the stationary distribution of the diffusion limit in Theorem 4.3. All these results are established under Assumptions 3.1, 3.2 and 3.3, and a further condition that  $\hat{Q}^n$  is stochastically bounded. At last, we show that  $\hat{Q}^n$  is stochastically bounded under Assumptions 3.1, 3.2, and 3.4, and consequently, all results in Theorem 4.1 – 4.3 hold (see Corollary 4.4 for details).

Recall the parameters  $\lambda, c$ , and  $\sigma_1, \sigma_{-1}$  in (3.4), (3.1), and Assumption 3.1, respectively, and the functions  $H_1, H_{-1}$  in Assumption 3.3.

**Theorem 4.1.** *Assume that  $\hat{Q}^n$  is stochastically bounded and  $\hat{Q}^n(0)$  converges weakly to some random variable  $q$  as  $n \rightarrow \infty$ . Then under Assumptions 3.1, 3.2 and 3.3, we have*

$$\left( \hat{W}_1^n - \frac{\hat{Q}^{n,+}}{\lambda}, \hat{W}_{-1}^n - \frac{\hat{Q}^{n,-}}{\lambda} \right) \Rightarrow 0, \quad \text{as } n \rightarrow \infty.$$

**Theorem 4.2.** *Under the same conditions as in Theorem 4.1, we have  $\hat{Q}^n \Rightarrow Q$  as  $n \rightarrow \infty$ , where for  $t \geq 0$ ,*

$$Q(t) = q + \sqrt{\lambda^3(\sigma_1^2 + \sigma_{-1}^2)} W(t) + ct - \lambda \int_0^t H_1\left(\frac{Q^+(s)}{\lambda}\right) ds + \lambda \int_0^t H_{-1}\left(\frac{Q^-(s)}{\lambda}\right) ds, \quad (4.1)$$

with  $W$  a standard Brownian motion independent of  $q$ .

**Theorem 4.3.** *Let  $Q$  be as in Theorem 4.2. Assume that*

$$\lim_{x \rightarrow \infty} H_1(x) > \frac{c}{\lambda}, \quad \text{if } c \geq 0,$$

and

$$\lim_{x \rightarrow \infty} H_{-1}(x) > -\frac{c}{\lambda}, \quad \text{if } c \leq 0.$$

Then  $Q$  has a unique stationary distribution with density function given by

$$\pi(dx) = \begin{cases} C_0 \exp \left\{ -\frac{2}{\lambda^3(\sigma_1^2 + \sigma_{-1}^2)} \left( -cx + \lambda^2 \int_0^{\frac{x}{\lambda}} H_1(u) du \right) \right\} dx, & \text{if } x \geq 0, \\ C_0 \exp \left\{ -\frac{2}{\lambda^3(\sigma_1^2 + \sigma_{-1}^2)} \left( -cx + \lambda^2 \int_0^{-\frac{x}{\lambda}} H_{-1}(u) du \right) \right\} dx, & \text{if } x < 0, \end{cases}$$

where  $C_0$  is a positive constant such that  $\int_{\mathbb{R}} \pi(dx) = 1$ .



**Corollary 4.4.** *Under Assumptions 3.1, 3.2 and 3.4,  $\hat{Q}^n$  is stochastically bounded, and if  $\hat{Q}^n(0)$  converges weakly to some random variable  $q$  as  $n \rightarrow \infty$ , then  $\hat{Q}^n \Rightarrow Q$  as  $n \rightarrow \infty$ , where*

$$Q(t) = q + \sqrt{\lambda^3(\sigma_1^2 + \sigma_{-1}^2)}W(t) + ct - \lambda \int_0^t \int_0^{\frac{Q^+(s)}{\lambda}} h_1(u) du ds + \lambda \int_0^t \int_0^{\frac{Q^-(s)}{\lambda}} h_{-1}(u) du ds, \quad t \geq 0,$$

with  $W$  a standard Brownian motion. Furthermore, assume that

$$\int_0^\infty h_1(u) du > \frac{c}{\lambda}, \quad \text{if } c \geq 0,$$

and

$$\int_0^\infty h_{-1}(u) du > -\frac{c}{\lambda}, \quad \text{if } c \leq 0.$$

Then  $Q$  has a unique stationary distribution with density function given by

$$\pi(dx) = \begin{cases} C_0 \exp \left\{ -\frac{2}{\lambda^3(\sigma_1^2 + \sigma_{-1}^2)} \left( -cx + \lambda \int_0^x \int_0^{\frac{s}{\lambda}} h_1(u) du ds \right) \right\} dx, & \text{if } x \geq 0, \\ C_0 \exp \left\{ -\frac{2}{\lambda^3(\sigma_1^2 + \sigma_{-1}^2)} \left( -cx + \lambda \int_0^{-x} \int_0^{\frac{s}{\lambda}} h_{-1}(u) du ds \right) \right\} dx, & \text{if } x < 0, \end{cases}$$

where  $C_0$  is a positive constant such that  $\int_{\mathbb{R}} \pi(dx) = 1$ .

## 5. Proofs

We first re-characterize the offered waiting times in the following lemma, which will be used in other proofs. Recall from (2.7), (2.8) and (2.10) that for  $i = \pm 1$  and  $k \in \mathbb{N}$ ,

$$w_{i,k}^n = \left( t_{-i,k+Q_i^n(0)-R_i^n(t_{i,k}^n-)-Q_{-i}^n(0)+R_{-i}^n(t_{i,k}^n-)} - t_{i,k}^n \right) 1_{\{Q_{-i}^n(t_{i,k}^n-)=0\}}.$$

Recall from Remark 2.1 that  $k - R_{-1}^n(t_{-1,k}^n-) - Q_1^n(0) + R_1^n(t_{-1,k}^n-)$  could be negative or zero.

**Lemma 5.1.** *For  $k \in \mathbb{N}$ , assume  $k - R_{-1}^n(t_{-1,k}^n-) - Q_1^n(0) + R_1^n(t_{-1,k}^n-) > 0$ . Then for  $i = \pm 1$ , when  $Q_{-i}^n(t_{i,k}^n-) = 0$ ,*

$$t_{-i,k+Q_i^n(0)-R_i^n(t_{i,k}^n-)-Q_{-i}^n(0)+R_{-i}^n(t_{i,k}^n-)} \geq t_{i,k}^n, \quad (5.1)$$

and when  $Q_{-i}^n(t_{i,k}^n-) > 0$ ,

$$t_{-i,k+Q_i^n(0)-R_i^n(t_{i,k}^n-)-Q_{-i}^n(0)+R_{-i}^n(t_{i,k}^n-)} \leq t_{i,k}^n. \quad (5.2)$$

If  $k - R_{-1}^n(t_{-1,k}^n-) - Q_1^n(0) + R_1^n(t_{-1,k}^n-) \leq 0$ , then  $Q_1^n(t_{-1,k}^n-) > 0$ , and  $w_{-1,k}^n = 0$ . Consequently, we have for  $i = \pm 1$ , and  $k \in \mathbb{N}$ ,

$$w_{i,k}^n = \left[ t_{-i,k+Q_i^n(0)-R_i^n(t_{i,k}^n-)-Q_{-i}^n(0)+R_{-i}^n(t_{i,k}^n-)} - t_{i,k}^n \right]^+. \quad (5.3)$$

*Proof.* Let  $i = \pm 1$ . For  $t \geq 0$ , when  $Q_i^n(t) = 0$ , we observe that  $R_i^n(t) = G_i^n(t)$ , and  $Q_i^n(0) + N_i^n(t) - R_i^n(t)$  is the number of the matched pairs that leave the system by time  $t$ . On the other hand, if  $Q_i^n(t) > 0$ , then  $Q_i^n(0) + N_i^n(t) - R_i^n(t)$  will be greater than or equal to the number of the matched pairs that leave the system by time  $t$ . So when  $Q_i^n(t) = 0$ , we have  $Q_{-i}^n(t) \geq 0$ , and

$$Q_i^n(0) + N_i^n(t) - R_i^n(t) \leq Q_{-i}^n(0) + N_{-i}^n(t) - R_{-i}^n(t). \quad (5.4)$$

Thus for  $k \in \mathbb{N}$ , if  $Q_{-i}^n(t_{i,k}^n-) = 0$ , we have

$$N_{-i}^n(t_{i,k}^n-) + Q_{-i}^n(0) - R_{-i}^n(t_{i,k}^n-) \leq N_i^n(t_{i,k}^n-) + Q_i^n(0) - R_i^n(t_{i,k}^n-) = k - 1 + Q_i^n(0) - R_i^n(t_{i,k}^n-).$$

Thus

$$k + Q_i^n(0) - R_i^n(t_{i,k}^n-) - Q_{-i}^n(0) + R_{-i}^n(t_{i,k}^n-) \geq N_{-i}^n(t_{i,k}^n-) + 1. \quad (5.5)$$

From (5.5), if  $k - R_{-1}^n(t_{-1,k}^n-) - Q_1^n(0) + R_1^n(t_{-1,k}^n-) \leq 0$ , then  $Q_1^n(t_{-1,k}^n-)$  cannot be zero, and must be positive. Now assume  $k - R_{-1}^n(t_{-1,k}^n-) - Q_1^n(0) + R_1^n(t_{-1,k}^n-) > 0$ . Again from (5.5), we have

$$t_{-i,k+Q_i^n(0)-R_i^n(t_{i,k}^n-)-Q_{-i}^n(0)+R_{-i}^n(t_{i,k}^n-)}^n \geq t_{-i,N_{-i}^n(t_{i,k}^n-)+1}^n \geq t_{i,k}^n.$$

This shows (5.1). When  $Q_{-i}^n(t_{i,k}^n-) > 0$ , we have  $Q_i^n(t_{i,k}^n-) = 0$ , which implies  $Q_i^n(t_{i,k}^n) = 0$  and  $G_i^n(t_{i,k}^n) = G_i^n(t_{i,k}^n-) = R_i^n(t_{i,k}^n-)$ . Similar to the above analysis,  $N_i^n(t_{i,k}^n) + Q_i^n(0) - G_i^n(t_{i,k}^n) = k + Q_i^n(0) - R_i^n(t_{i,k}^n-)$  is the number of matched pairs that leave the system by time  $t_{i,k}^n$ . Then we have

$$k + Q_i^n(0) - R_i^n(t_{i,k}^n-) \leq N_{-i}^n(t_{i,k}^n) + Q_{-i}^n(0) - R_{-i}^n(t_{i,k}^n) \leq N_{-i}^n(t_{i,k}^n) + Q_{-i}^n(0) - R_{-i}^n(t_{i,k}^n-).$$

Thus

$$k + Q_i^n(0) - R_i^n(t_{i,k}^n-) - Q_{-i}^n(0) + R_{-i}^n(t_{i,k}^n-) \leq N_{-i}^n(t_{i,k}^n),$$

and

$$t_{-i,k+Q_i^n(0)-R_i^n(t_{i,k}^n-)-Q_{-i}^n(0)+R_{-i}^n(t_{i,k}^n-)}^n \leq t_{-i,N_{-i}^n(t_{i,k}^n)}^n \leq t_{i,k}^n.$$

This shows (5.2). Finally, (5.3) follows immediately from (5.1) and (5.2).  $\square$

Let  $i = \pm 1$ . We next introduce the following filtrations. First recall that we assume  $Q_1(0) \geq 0$  and  $Q_{-1}(0) = 0$ . We will let  $w_{1,k}^n = 0$  when  $k < -Q_1^n(0) + 1$ , and  $w_{-1,k}^n = 0$  when  $k \leq 0$ . Now define for  $k = 0, -1, -2, \dots$ ,

$$\begin{aligned} \mathcal{F}_{1,k}^n &= \sigma \{ Q^n(0); (d_{1,l}^n, w_{1,l}^n), l = k+1, \dots, 0; R_{1,k}^n(0); \\ &\quad t_{-1,j}^n, j = 1, \dots, -k+1 - R_{1,k}^n(0) \}, \\ \mathcal{F}_{-1,k}^n &= \sigma \{ Q^n(0) \}, \end{aligned} \quad (5.6)$$

and for  $k \in \mathbb{N}$ ,

$$\begin{aligned} \mathcal{F}_{i,k}^n &= \sigma \{ Q^n(0); t_{i,l}^n, l = 1, 2, \dots, k; N_{-i}^n(s), 0 \leq s \leq t_{i,k}^n; \\ &\quad (w_{i,l}^n, d_{i,l}^n), l = -Q_i^n(0) + 1, \dots, 0, 1, \dots, k-1; \\ &\quad (w_{-i,l}^n, d_{-i,l}^n), l = -Q_{-i}^n(0) + 1, \dots, 0, 1, \dots, N_{-i}^n(t_{i,k}^n-); \\ &\quad t_{-i,l}^n, l = 1, 2, \dots, k + Q_i^n(0) + R_i^n(t_{i,k}^n-) - Q_{-i}^n(0) + R_{-i}^n(t_{i,k}^n-) \}. \end{aligned} \quad (5.7)$$

**Lemma 5.2.** For  $k \in \mathbb{Z}$ ,  $w_{i,k}^n \in \mathcal{F}_{i,k}^n$ , and  $d_{i,k}^n$  is independent of  $\mathcal{F}_{i,k}^n$ .

*Proof.* From (2.7) and (2.8), it is clear that  $w_{i,k}^n \in \mathcal{F}_{i,k}^n, k \in \mathbb{Z}$ . To see  $d_{i,k}^n$  is independent of  $\mathcal{F}_{i,k}^n$ , it suffices to show that  $d_{i,k}^n$  is independent of the offered waiting times of all customers from both classes who arrive before  $t_{i,k}^n$ , because we assume  $Q^n(0), \{u_{1,k}^n : k \in \mathbb{N}\}, \{u_{-1,k}^n : k \in \mathbb{N}\}, \{d_{1,k}^n : k \in \mathbb{Z}\}$ , and  $\{d_{-1,k}^n : k \in \mathbb{Z}\}$  are independent. Fix  $k \in \mathbb{Z}$ . We will use mathematical induction in the following. For the first customer who arrives at the system (prior to time 0 or after time 0), assuming he/she is from Class  $i$ , the offered waiting time is  $t_{-i,1}$ , which is independent of  $d_{i,k}$ . Assume that  $d_{i,k}$  is independent of the offered waiting times of the first  $l$  customers (among all customers of both classes) who arrive at the system before  $t_{i,k}^n$ . We next note that for any  $l \in \mathbb{Z}$  and  $i = \pm 1$ , the offered waiting time  $w_{i,l}$  of the  $(l1_{\{l>0\}} + (-l+1)1_{\{l \leq 0\}})$ -th customer of Class  $i$  is determined by the offered waiting times of all customers from both classes who arrive before him/her, and other random variables that are independent of  $d_{i,k}$ . Thus  $d_{i,k}$  is also independent of the offered waiting time of the  $((l+1)1_{\{l>0\}} + (-l+2)1_{\{l \leq 0\}})$ -th customer (among all customers of both classes) who arrive at the system before  $t_{i,k}^n$ . The lemma now follows.  $\square$

We divide the rest of the section into four subsections. In Section 5.1, we decompose  $\hat{W}_i^n$  into several processes, and study the asymptotic behaviors of these processes (see Lemmas 5.5, 5.7, 5.8). In Section 5.2, using the results from the previous subsection, we prove the  $C$ -tightness of  $(\hat{W}_1^n, \hat{W}_{-1}^n)$  in Lemma 5.9, and establish the weak convergence of  $(\hat{W}_1^n, \hat{W}_{-1}^n)$  in Theorem 5.11. The weak limit is uniquely characterized by a continuous functional defined in Proposition 5.10. Section 5.3 is then devoted to derive the weak convergence of  $\hat{R}_i^n$  and  $\hat{G}_i^n$ , and as shown in Lemmas 5.12 and 5.13, they have the same weak limit that can be characterized in terms of the offered waiting time processes. We also provide all the proofs of Theorems 4.1, 4.2, and 4.3 in this section. Finally, the proof of Corollary 4.4 can be found in Section 5.4.

### 5.1. Decomposition of $\hat{W}_i^n$

We decompose  $\hat{W}_i^n$  into several stochastic processes, and these processes will be analyzed separately in Lemmas 5.5, 5.7, 5.8.

**Lemma 5.3.** For  $i = \pm 1$  and  $t \geq 0$ ,

$$\begin{aligned} \hat{W}_i^n(t) = & \left[ \hat{M}_{i,1}^n(t) + \hat{M}_{i,2}^n(t) - \frac{n}{\lambda_{-i}^n} \hat{M}_{i,3}^n(\bar{N}_i^n(t)) + \frac{n}{\lambda_{-i}^n} \hat{M}_{-i,3}^n(\bar{N}_{-i}^n(t)) \right. \\ & - \frac{\sqrt{n}}{\lambda_{-i}^n} \int_0^t F_i^n(W_i^n(u-)) dN_i^n(u) + \frac{\sqrt{n}}{\lambda_{-i}^n} \int_0^t F_{-i}^n(W_{-i}^n(u-)) dN_{-i}^n(u) \\ & \left. + \frac{n}{\lambda_{-i}^n} \left( \frac{1}{\sqrt{n}} + \hat{Q}_i^n(0) - \hat{Q}_{-i}^n(0) - \hat{\xi}_i^n + \hat{\xi}_{-i}^n \right) \right]^+, \end{aligned} \quad (5.8)$$

where

$$\hat{M}_{i,1}^n(t) = \sqrt{n} \left[ t_{-i, N_i^n(t)+1+Q_i^n(0)-R_i^n(t)-Q_{-i}^n(0)+R_{-i}^n(t)}^n - \frac{1}{\lambda_{-i}^n} (N_i^n(t) + 1 + Q_i^n(0) - R_i^n(t) - Q_{-i}^n(0) + R_{-i}^n(t)) \right], \quad (5.9)$$

$$\hat{M}_{i,2}^n(t) = \frac{\sqrt{n}}{\lambda_{-i}^n} (N_i^n(t) - \lambda_{-i}^n t), \quad (5.10)$$

$$\hat{M}_{i,3}^n(t) = \frac{1}{\sqrt{n}} \sum_{k=1}^{[nt]} \left( 1_{\{d_{i,k}^n < w_{i,k}^n\}} - F_i^n(w_{i,k}^n) \right), \quad (5.11)$$

$$\hat{\xi}_i^n = \frac{1}{\sqrt{n}} \sum_{k=-Q_i^n(0)+1}^0 1_{\{d_{i,k}^n < w_{i,k}^n\}}. \quad (5.12)$$

*Proof.* It suffices to show that

$$\hat{R}_i^n(t) = \hat{\xi}_i^n + \hat{M}_{i,3}^n(\bar{N}_i^n(t)) + \frac{1}{\sqrt{n}} \int_0^t F_i^n(W_i^n(u-)) dN_i^n(u), \quad t \geq 0. \quad (5.13)$$

From Lemma 5.2, for  $k \in \mathbb{N}$ ,  $w_{i,k}^n \in \mathcal{F}_{i,k}^n$ , and  $d_{i,k}^n$  is independent of  $\mathcal{F}_{i,k}^n$ . Thus for  $t \geq 0$ ,

$$\begin{aligned} \hat{R}_i^n(t) &= \frac{1}{\sqrt{n}} \sum_{k=-Q_i^n(0)+1}^{N_i^n(t)} 1_{\{d_{i,k}^n < w_{i,k}^n\}} \\ &= \frac{1}{\sqrt{n}} \sum_{k=-Q_i^n(0)+1}^0 1_{\{d_{i,k}^n < w_{i,k}^n\}} + \frac{1}{\sqrt{n}} \sum_{k=1}^{N_i^n(t)} 1_{\{d_{i,k}^n < w_{i,k}^n\}} \\ &= \hat{\xi}_i^n + \frac{1}{\sqrt{n}} \sum_{k=1}^{N_i^n(t)} \left( 1_{\{d_{i,k}^n < w_{i,k}^n\}} - \mathbb{E} \left( 1_{\{d_{i,k}^n < w_{i,k}^n\}} | \mathcal{F}_{i,k}^n \right) \right) + \frac{1}{\sqrt{n}} \sum_{k=1}^{N_i^n(t)} \mathbb{E} \left( 1_{\{d_{i,k}^n < w_{i,k}^n\}} | \mathcal{F}_{i,k}^n \right) \\ &= \hat{\xi}_i^n + \frac{1}{\sqrt{n}} \sum_{k=1}^{N_i^n(t)} \left( 1_{\{d_{i,k}^n < w_{i,k}^n\}} - F_i^n(w_{i,k}^n) \right) + \frac{1}{\sqrt{n}} \sum_{k=1}^{N_i^n(t)} F_i^n(w_{i,k}^n) \\ &= \hat{\xi}_i^n + \frac{1}{\sqrt{n}} \sum_{k=1}^{N_i^n(t)} \left( 1_{\{d_{i,k}^n < w_{i,k}^n\}} - F_i^n(w_{i,k}^n) \right) + \frac{1}{\sqrt{n}} \int_0^t F_i^n(W_i^n(u-)) dN_i^n(u). \end{aligned}$$

The lemma follows.  $\square$

We first establish the fluid limit of the state process in Lemma 5.4, and then study the processes in (5.9) – (5.12) as  $n \rightarrow \infty$  in Lemmas 5.5, 5.7, 5.8.

**Lemma 5.4.** *Assume that  $\hat{Q}^n$  is stochastically bounded. Then under Assumptions 3.1 and 3.3, we have*

$$(\bar{N}_1^n, \bar{N}_{-1}^n, \bar{R}_1^n, \bar{R}_{-1}^n, W_1^n, W_{-1}^n, \bar{Q}^n) \Rightarrow (\lambda_1, \lambda_{-1}, 0, 0, 0, 0, 0)\iota, \quad (5.14)$$

where  $\iota : [0, \infty) \rightarrow [0, \infty)$  is the identity map.

*Proof.* Using functional law of large numbers for renewal processes (see Theorem 1 in [13]), we have

$$(\bar{N}_1^n, \bar{N}_{-1}^n) \Rightarrow (\lambda_1, \lambda_{-1})_t. \quad (5.15)$$

Recall that for  $t \geq 0$ ,

$$W_i^n(t) = \left( t_{-i, N_i^n(t)+1+Q_i^n(0)-R_i^n(t)-Q_{-i}^n(0)+R_{-i}^n(t)} - t \right) 1_{\{Q_{-i}^n(t)=0\}}.$$

When  $Q_{-i}^n(t) = 0$ , we have  $R_{-i}^n(t) = G_{-i}^n(t)$ . Thus

$$W_i^n(t) \leq |t_{-i, N_i^n(t)+1+Q_i^n(0)-G_i^n(t)-Q_{-i}^n(0)+G_{-i}^n(t)} - t|.$$

We first consider  $W_1^n$ , and note that from (2.2), for  $t \geq 0$ ,

$$|N_1^n(t) + 1 + Q_1^n(0) - G_1^n(t) - Q_{-1}^n(0) + G_{-1}^n(t)| = |Q^n(t) + 1 + N_{-1}^n(t)| \leq |Q^n(t)| + 1 + N_{-1}^n(t).$$

Consequently, we have for  $T \geq 0$ ,

$$\begin{aligned} \sup_{0 \leq t \leq T} W_1^n(t) &\leq \sup_{0 \leq t \leq T} \left| \sum_{k=1}^{|Q^n(t)|+1+N_{-1}^n(t)} u_{-1,k}^n - t \right| \\ &= \sup_{0 \leq t \leq T} \left| \sum_{k=1}^{|Q^n(t)|+1+N_{-1}^n(t)} \left( u_{-1,k}^n - \frac{1}{\lambda_{-1}^n} \right) + \frac{1}{\lambda_{-1}^n} (|Q^n(t)| + 1) + \frac{1}{\lambda_{-1}^n} (N_{-1}^n(t) - \lambda_{-1}^n t) \right| \\ &\leq \sup_{0 \leq t \leq T} \left| \sum_{k=1}^{|Q^n(t)|+1+N_{-1}^n(t)} \left( u_{-1,k}^n - \frac{1}{\lambda_{-1}^n} \right) \right| + \frac{n}{\lambda_{-1}^n} \|\bar{Q}^n\|_t + \frac{1}{\lambda_{-1}^n} + \frac{n}{\lambda_{-1}^n} \frac{\|\hat{N}_{-1}^n\|_t}{\sqrt{n}}. \end{aligned} \quad (5.16)$$

Using functional central limit theorems for triangular arrays and renewal processes (see again Theorem 1 in [13]), under Assumption 3.1,

$$\left( \sqrt{n} \sum_{k=1}^{\lfloor n \cdot \rfloor} \left( u_{-1,k}^n - \frac{1}{\lambda_{-1}^n} \right), \hat{N}_{-1}^n \right) \Rightarrow (\sigma_{-1} B, -\sigma_{-1} \lambda_{-1}^{3/2} B), \quad (5.17)$$

where  $B$  is a standard Brownian motion. Noting that  $\bar{Q}^n \Rightarrow 0$ , and applying the random time change theorem to the first term in (5.16), we have  $W_1 \Rightarrow 0$ . Similarly, we can show the same convergence result for  $W_{-1}^n$ . Thus

$$W_i^n \Rightarrow 0. \quad (5.18)$$

Next from (5.13), for  $t \geq 0$ ,

$$\bar{R}_i^n(t) = \frac{1}{\sqrt{n}} \hat{\xi}_i^n + \frac{1}{\sqrt{n}} \hat{M}_{i,3}^n(\bar{N}_i^n(t)) + \frac{1}{n} \int_0^t F_i^n(W_i^n(u-)) dN_i^n(u), \quad t \geq 0.$$

It is clear that

$$\frac{1}{\sqrt{n}} \hat{\xi}_i^n \leq |\bar{Q}^n(0)| \Rightarrow 0. \quad (5.19)$$

We next consider  $\frac{1}{\sqrt{n}}\hat{M}_{i,3}^n$  and note that  $\{\frac{1}{\sqrt{n}}\hat{M}_{i,3}^n(t)\}_{t \geq 0}$  is a  $\{\mathcal{F}_{i,t}^n \vee \sigma(d_{i,[nt]}^n)\}_{t \geq 0}$  martingale, where  $\{\mathcal{F}_{i,k}^n\}_{k \in \mathbb{N}}$  is defined in (5.7). From Doob's inequality for martingales, we have for  $T \geq 0$ ,

$$\mathbb{E} \left( \sup_{0 \leq t \leq T} \left( \frac{1}{\sqrt{n}} \hat{M}_{i,3}^n(t) \right)^2 \right) \leq 4\mathbb{E} \left( \left( \frac{1}{\sqrt{n}} \hat{M}_{i,3}^n(T) \right)^2 \right) = 4\mathbb{E} \left( \left[ \frac{1}{\sqrt{n}} \hat{M}_{i,3}^n \right](T) \right) \leq \frac{8nT}{n^2} \rightarrow 0.$$

From the random time change theorem, we have

$$\frac{1}{\sqrt{n}} \hat{M}_{i,3}^n(\bar{N}_i^n(\cdot)) \Rightarrow 0. \quad (5.20)$$

Next using Assumption 3.3, we have that for  $\delta > 0$ ,

$$\begin{aligned} & \limsup_{n \rightarrow \infty} \mathbb{P} \left( \frac{1}{n} \int_0^t F_i^n(W_i^n(u-)) dN_i^n(u) > \delta \right) \\ &= \lim_{K \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P} \left( \frac{1}{n} \int_0^t F_i^n(W_i^n(u-)) dN_i^n(u) > \delta, \|\hat{W}_i^n\|_t \leq K \right) + \lim_{K \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P}(\|\hat{W}_i^n\|_t > K). \end{aligned} \quad (5.21)$$

From (5.16), (5.17), and the assumption that  $\hat{Q}^n$  is stochastically bounded, we see that  $\hat{W}_i^n$  is also stochastically bounded, and so

$$\lim_{K \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P}(\|\hat{W}_i^n\|_t > K) = 0.$$

Therefore, (5.21) can be bounded by

$$\lim_{K \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P} \left( \bar{N}_i^n(t) > \frac{\delta}{F_i^n(K/\sqrt{n})} \right) = 0.$$

This shows that

$$\frac{1}{n} \int_0^\cdot F_i^n(W_i^n(u-)) dN_i^n(u) \Rightarrow 0. \quad (5.22)$$

Combining (5.19), (5.20), and (5.22), we have

$$\bar{R}_i^n \Rightarrow 0. \quad (5.23)$$

The lemma follows from (5.15), (5.18), and (5.23).  $\square$

**Lemma 5.5.** Assume that  $\hat{Q}^n$  is stochastically bounded. Let  $\hat{M}_{i,1}^n$  and  $\hat{M}_{i,2}^n$  be as in (5.9) and (5.10). Then under Assumptions 3.1, 3.2 and 3.3,

$$(\hat{M}_{1,1}^n, \hat{M}_{1,2}^n, \hat{M}_{-1,1}^n, \hat{M}_{-1,2}^n) \Rightarrow (0, \frac{c}{\lambda}, 0, -\frac{c}{\lambda})\iota + (Y_{-1}, -Y_1, Y_1, -Y_{-1}),$$

where  $\iota : [0, \infty) \rightarrow [0, \infty)$  is the identity map, and  $Y_1(t) = \sigma_1 B_1(\lambda t)$  and  $Y_{-1}(t) = \sigma_{-1} B_{-1}(\lambda t)$  with  $B_1$  and  $B_{-1}$  being two independent standard Brownian motions.

*Proof.* Define

$$\tilde{M}_i^n(t) = \sqrt{n} \sum_{k=1}^{\lfloor \lambda_i^n t \rfloor} \left( u_{i,k}^n - \frac{1}{\lambda_i^n} \right), \quad t \geq 0. \quad (5.24)$$

Using Theorem 1 in [13], under Assumptions 3.1 and 3.2,

$$(\tilde{M}_1^n, \tilde{M}_{-1}^n, \frac{1}{\lambda} \hat{N}_1^n, \frac{1}{\lambda} \hat{N}_{-1}^n) \Rightarrow (Y_1, Y_{-1}, -Y_1, -Y_{-1}), \quad (5.25)$$

where  $Y_1(t) = \sigma_1 B_1(\lambda t)$  and  $Y_{-1}(t) = \sigma_{-1} B_{-1}(\lambda t)$  with  $B_1$  and  $B_{-1}$  being two independent standard Brownian motions. We next note that for  $i = \pm 1$  and  $t \geq 0$ ,

$$\hat{M}_{i,1}^n(t) = \tilde{M}_{-i}^n \left( \frac{n}{\lambda_{-i}^n} (\bar{N}_i^n(t) + \frac{1}{n} + \bar{Q}_i^n(0) - \bar{R}_i^n(t) - \bar{Q}_{-i}^n(0) + \bar{R}_{-i}^n(t)) \right), \quad (5.26)$$

and

$$\hat{M}_{i,2}^n(t) = \frac{n}{\lambda_{-i}^n} \hat{N}_i^n(t) + \frac{n}{\lambda_{-i}^n} \frac{(\lambda_i^n - \lambda_{-i}^n)t}{\sqrt{n}}. \quad (5.27)$$

The lemma follows by applying the random time change theorem, Lemma 5.4, and Assumptions 3.1 and 3.2 to  $(\hat{M}_{1,1}^n, \hat{M}_{1,2}^n, \hat{M}_{-1,1}^n, \hat{M}_{-1,2}^n)$ .  $\square$

The following lemma on martingale convergence (see [22] or [7]) will be used in the proof of Lemma 5.7.

**Lemma 5.6.** *Suppose  $\{M^n(t); t \geq 0\}$  is a local martingale w.r.t some filtrations, and for  $t \geq 0$ ,*

$$\mathbb{E} \left( \sup_{0 \leq s \leq t} |M^n(s) - M^n(s-)| \right) \rightarrow 0, \text{ and } [M^n]_t \Rightarrow 0, \text{ as } n \rightarrow \infty.$$

*Then  $M^n \Rightarrow 0$  as  $n \rightarrow \infty$ .*

**Lemma 5.7.** *Assume that  $\hat{Q}^n$  is stochastically bounded. Let  $\hat{M}_{i,3}^n$  be as in (5.11). Then under Assumption 3.1, 3.2, and 3.3,  $\hat{M}_{i,3}^n \Rightarrow 0$ .*

*Proof.* Recall that  $\{\hat{M}_{i,3}^n(t)\}_{t \geq 0}$  is an  $\{\mathcal{F}_{i,t}^n \vee \sigma(d_{i,[nt]}^n)\}_{t \geq 0}$  martingale, where  $\mathcal{F}_{i,t}^n$  is defined in (5.7), and its quadratic variation is

$$[\hat{M}_{i,3}^n]_t = \frac{1}{n} \sum_{k=1}^{[nt]} \left( 1_{\{w_{i,k}^n \geq d_{i,k}^n\}} - F_i^n(w_{i,k}^n) \right)^2.$$

We next observe that for  $t \geq 0$ ,

$$\begin{aligned} \mathbb{E}([\hat{M}_{i,3}^n]_t) &= \frac{1}{n} \sum_{k=1}^{[nt]} \mathbb{E} \left( 1_{\{w_{i,k}^n \geq d_{i,k}^n\}} - F_i^n(w_{i,k}^n) \right)^2 \\ &= \frac{1}{n} \sum_{k=1}^{[nt]} \mathbb{E} \left( 1_{\{w_{i,k}^n \geq d_{i,k}^n\}} - \mathbb{E}(1_{\{w_{i,k}^n \geq d_{i,k}^n\}} | \mathcal{F}_{i,k}^n) \right)^2 \\ &= \frac{1}{n} \sum_{k=1}^{[nt]} \left[ \mathbb{E} \left( 1_{\{w_{i,k}^n \geq d_{i,k}^n\}} \right) - \mathbb{E} \left( \mathbb{E}^2(1_{\{w_{i,k}^n \geq d_{i,k}^n\}} | \mathcal{F}_{i,k}^n) \right) \right] \\ &\leq \frac{1}{n} \sum_{k=1}^{[nt]} \mathbb{E} \left( 1_{\{w_{i,k}^n \geq d_{i,k}^n\}} \right) \end{aligned}$$

$$\begin{aligned}
& \stackrel{\text{/Double-ended queues with reneging}}{=} \mathbb{E} \left( \frac{1}{n} \sum_{k=1}^{\lfloor nt \rfloor} 1_{\{w_{i,k}^n \geq d_{i,k}^n\}} 1_{\{N_i^n(2t/\lambda) \geq nt\}} \right) + \mathbb{E} \left( \frac{1}{n} \sum_{k=1}^{\lfloor nt \rfloor} 1_{\{w_{i,k}^n \geq d_{i,k}^n\}} 1_{\{N_i^n(2t/\lambda) < nt\}} \right) \\
& \leq \mathbb{E} \left( \frac{1}{n} \sum_{k=1}^{N_i^n(2t/\lambda)} 1_{\{w_{i,k}^n \geq d_{i,k}^n\}} \right) + t\mathbb{P}(N_i^n(2t/\lambda) < nt) \\
& = \mathbb{E}(\bar{R}_i^n(2t/\lambda)) + t\mathbb{P}(\bar{N}_i^n(2t/\lambda) < t).
\end{aligned} \tag{16}$$

Noting that  $\bar{N}_i^n(2t/\lambda) \Rightarrow 2t$  as  $n \rightarrow \infty$ , thus  $\mathbb{P}(\bar{N}_i^n(2t/\lambda) < t) \rightarrow 0$ . Next from Lemma 5.4, we have  $\bar{R}_i^n \Rightarrow 0$ , and we further observe that  $\bar{R}_i^n$  is uniformly integrable (which follows from the uniform integrability of  $\bar{N}_i^n$ ). Thus  $\mathbb{E}(\bar{R}_i^n(2t/\lambda)) \rightarrow 0$ . Finally, the result follows from Lemma 5.6.  $\square$

**Lemma 5.8.** Assume  $\hat{Q}^n(0)$  is stochastically bounded. Let  $\hat{\xi}_i$  be as in (5.12). Then under Assumptions 3.1, 3.3, we have  $\hat{\xi}_i^n \Rightarrow 0$ .

*Proof.* Noting that we assume  $Q_1(0) \geq 0$  and  $Q_{-1}(0) = 0$ , so  $\hat{\xi}_{-1}^n \equiv 0$ , and it only needs to show  $\hat{\xi}_1^n \Rightarrow 0$ . Recall  $\{\mathcal{F}_{1,k}^n : k = 0, -1, -2, \dots\}$  defined in (5.6). From Lemma 5.2,  $w_{1,k}^n \in \mathcal{F}_{1,k}^n$  and  $d_{1,k}^n$  is independent of  $\mathcal{F}_{1,k}^n$ ,  $k \in \mathbb{Z}$ . Define for  $t \geq 0$ ,

$$\tilde{M}^n(t) = \frac{1}{\sqrt{n}} \sum_{k=-\lfloor nt \rfloor + 1}^0 [1_{\{d_{1,k}^n \leq w_{1,k}^n\}} - F_1^n(w_{1,k}^n)]$$

We see that  $\{\tilde{M}^n(t)\}_{t \geq 0}$  is a  $\{\mathcal{F}_{1,-\lfloor nt \rfloor}^n \vee \sigma(d_{1,-\lfloor nt \rfloor}^n)\}_{t \geq 0}$  martingale, and

$$[\tilde{M}^n]_t = \frac{1}{n} \sum_{k=-\lfloor nt \rfloor + 1}^0 [1_{\{d_{1,k}^n \leq w_{1,k}^n\}} - F_1^n(w_{1,k}^n)]^2, \quad t \geq 0.$$

It is clear that  $\{[\tilde{M}^n]_t\}_{t \geq 0}$  is  $C$ -tight and from Theorem VI.4.13 of [14], we conclude that  $\{\tilde{M}^n(t)\}_{t \geq 0}$  is also  $C$ -tight. Using the fact that  $\bar{Q}^n(0) \Rightarrow 0$ , we obtain that

$$\tilde{M}^n(\bar{Q}^n(0)) \Rightarrow 0, \quad \text{as } n \rightarrow \infty. \tag{5.28}$$

Thus it suffices to show that for  $\delta > 0$ ,

$$\lim_{n \rightarrow \infty} \mathbb{P} \left( \frac{1}{\sqrt{n}} \sum_{k=-Q^n(0)+1}^0 F_1^n(w_{1,k}^n) > \delta \right) = 0.$$

We observe that for all  $k = -Q^n(0) + 1, \dots, -1, 0$ ,

$$\sqrt{n}w_{1,k}^n \leq \sqrt{nt}_{-1, Q^n(0)}^n = \sqrt{n} \sum_{l=1}^{Q^n(0)} u_{-1,l}^n = \tilde{M}_{-1}^n(Q^n(0)/\lambda_{-1}^n) + \frac{n}{\lambda_{-1}^n} \hat{Q}^n(0),$$



where  $\tilde{M}_{-1}^n$  is defined in (5.24). Thus for each  $k = -Q^n(0) + 1, \dots, -1, 0$ ,  $\sqrt{n}w_{1,k}^n$  is stochastically bounded, and for  $\delta > 0$ ,

$$\begin{aligned}
& \limsup_{n \rightarrow \infty} \mathbb{P} \left( \frac{1}{\sqrt{n}} \sum_{k=-Q^n(0)+1}^0 F_1^n(w_{1,k}^n) > \delta \right) \\
& \leq \lim_{K \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P} \left( \frac{1}{\sqrt{n}} \sum_{k=-Q^n(0)+1}^0 F_1^n(w_{1,k}^n) > \delta, \sqrt{n}w_{1,k}^n \leq K \right) \\
& \quad + \lim_{K \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P}(\sqrt{n}w_{1,k}^n > K) \\
& \leq \lim_{K \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P}(\sqrt{n}F_1^n(K/\sqrt{n})\bar{Q}^n(0) > \delta) \\
& \quad + \lim_{K \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P}(\sqrt{n}w_{1,k}^n > K) \\
& = 0.
\end{aligned}$$

Combining the above convergence and (5.28), we have  $\hat{\xi}_1^n \Rightarrow 0$ .  $\square$

## 5.2. Weak convergence of $(\hat{W}_1^n, \hat{W}_{-1}^n)$

We prove the  $C$ -tightness of  $(\hat{W}_1^n, \hat{W}_{-1}^n)$  in Lemma 5.9, and establish the weak convergence of  $(\hat{W}_1^n, \hat{W}_{-1}^n)$  in Theorem 5.11. The weak limit is uniquely characterized by a continuous functional defined in Proposition 5.10.

**Lemma 5.9.** *Assume that  $\hat{Q}^n$  is stochastically bounded. Then under Assumptions 3.1, 3.2 and 3.3,  $(\hat{W}_1^n, \hat{W}_{-1}^n)$  is  $C$ -tight.*

*Proof.* From (5.16) and (5.17) in the proof of Lemma 5.4, we know that  $\hat{W}_i^n$  is stochastically bounded. Now from (5.8), for  $0 \leq s \leq t < \infty$ ,

$$\begin{aligned}
|\hat{W}_i^n(t) - \hat{W}_i^n(s)| & \leq |\hat{M}_{i,1}^n(t) - \hat{M}_{i,1}^n(s)| + |\hat{M}_{i,2}^n(t) - \hat{M}_{i,2}^n(s)| + \frac{n}{\lambda_{-i}^n} |\hat{M}_{i,3}^n(\bar{N}_i^n(t)) - \hat{M}_{i,3}^n(\bar{N}_i^n(s))| \\
& \quad + \frac{n}{\lambda_{-i}^n} |\hat{M}_{-i,3}^n(\bar{N}_{-i}^n(t)) - \hat{M}_{-i,3}^n(\bar{N}_{-i}^n(s))| \\
& \quad + \frac{\sqrt{n}}{\lambda_{-i}^n} \int_s^t F_i^n(W_i^n(u-)) dN_i^n(u) + \frac{\sqrt{n}}{\lambda_{-i}^n} \int_s^t F_{-i}^n(W_{-i}^n(u-)) dN_{-i}^n(u).
\end{aligned}$$

From Lemmas 5.4, 5.5 and 5.7,  $(\hat{M}_{i,1}^n(\cdot), \hat{M}_{i,2}^n(\cdot), \hat{M}_{i,3}^n(\bar{N}_i^n(\cdot)), \hat{M}_{-i,3}^n(\bar{N}_{-i}^n(\cdot)))$  are weakly convergent. We next note that for  $\delta > 0$ ,

$$\begin{aligned}
& \mathbb{P} \left( \frac{\sqrt{n}}{\lambda_{-i}^n} \int_s^t F_i^n(W_i^n(u-)) dN_i^n(u) > \delta \right) \\
& = \mathbb{P} \left( \frac{1}{\lambda_{-i}^n} \int_s^t \sqrt{n} F_i^n(W_i^n(u-)) dN_i^n(u) > \delta, \|\hat{W}_i^n\|_{s,t} \leq K \right) + \mathbb{P}(\|\hat{W}_i^n\|_{s,t} > K) \\
& \leq \mathbb{P} \left( \sqrt{n} F_i^n(K/\sqrt{n}) \frac{n}{\lambda_{-i}^n} (\bar{N}_i^n(t) - \bar{N}_i^n(s)) > \delta \right) + \mathbb{P}(\|\hat{W}_i^n\|_t > K).
\end{aligned}$$

Using Assumption 3.3 and noting that  $\bar{N}_i^n$  is  $C$ -tight and  $\bar{W}_i^n$  is stochastically bounded, we have that for  $\delta > 0$ ,

$$\begin{aligned} & \lim_{(t-s) \downarrow 0} \limsup_{n \rightarrow \infty} \mathbb{P} \left( \frac{\sqrt{n}}{\lambda_{-i}^n} \int_s^t F_i^n(W_i^n(u-)) dN_i^n(u) > \delta \right) \\ & \leq \lim_{K \rightarrow \infty} \lim_{(t-s) \downarrow 0} \limsup_{n \rightarrow \infty} \mathbb{P} \left( \sqrt{n} F_i^n(K/\sqrt{n}) \frac{n}{\lambda_{-i}^n} (\bar{N}_i^n(t) - \bar{N}_i^n(s)) > \delta \right) \\ & \quad + \lim_{K \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P}(\|\hat{W}_i^n\|_t > K) \\ & = 0. \end{aligned}$$

The result follows.  $\square$

**Proposition 5.10.** *Let  $H_1, H_{-1} : [0, \infty) \rightarrow [0, \infty)$  be nonnegative locally Lipschitz continuous functions as in Assumption 3.3.*

- (i) *Given  $x \in D([0, \infty), \mathbb{R})$ , there exists a unique pair of  $(w_1, w_{-1})$  such that  $w_i \in D([0, \infty), \mathbb{R}_+)$ ,  $i = \pm 1$ , and for  $t \geq 0$ ,*

$$w_1(t) = \left[ x(t) - \int_0^t H_1(w_1(s)) ds + \int_0^t H_{-1}(w_{-1}(s)) ds \right]^+, \quad (5.29)$$

$$w_{-1}(t) = \left[ x(t) - \int_0^t H_1(w_1(s)) ds + \int_0^t H_{-1}(w_{-1}(s)) ds \right]^-. \quad (5.30)$$

- (ii) *Define the functionals  $\Psi_1, \Psi_{-1} : D([0, \infty); \mathbb{R}) \rightarrow D([0, \infty); \mathbb{R}_+)$  by*

$$(\Psi_1, \Psi_{-1})(x) = (w_1, w_{-1}). \quad (5.31)$$

*Then  $(\Psi_1, \Psi_{-1})$  is continuous on  $D([0, \infty); \mathbb{R})$  with Skorohod  $J_1$ -topology.*

*Proof.* See Appendix.  $\square$

**Theorem 5.11.** *Assume that  $\hat{Q}^n(0)$  converges weakly to some random variable  $q$ , and  $\hat{Q}^n$  is stochastically bounded. Then under Assumptions 3.1, 3.2, and 3.3,*

$$(\hat{W}_1^n, \hat{W}_{-1}^n) \Rightarrow (\Psi_1, \Psi_{-1})(X),$$

*where  $X$  is a Brownian motion with drift  $\frac{c}{\lambda}$  and variance  $\lambda(\sigma_1^2 + \sigma_{-1}^2)$ , and initial value  $X(0) = \frac{q}{\lambda}$ .*

*Proof.* From (5.8), for  $i = \pm 1$  and  $t \geq 0$ ,

$$\begin{aligned} \hat{W}_i^n(t) = & \left[ \hat{M}_{i,1}^n(t) + \hat{M}_{i,2}^n(t) - \frac{n}{\lambda_{-i}^n} \hat{M}_{i,3}^n(\bar{N}_i^n(t)) + \frac{n}{\lambda_{-i}^n} \hat{M}_{-i,3}^n(\bar{N}_{-i}^n(t)) \right. \\ & - \frac{\sqrt{n}}{\lambda_{-i}^n} \int_0^t F_i^n(W_i^n(u-)) dN_i^n(u) + \frac{\sqrt{n}}{\lambda_{-i}^n} \int_0^t F_{-i}^n(W_{-i}^n(u-)) dN_{-i}^n(u) \\ & \left. + \frac{n}{\lambda_{-i}^n} \left( \frac{1}{\sqrt{n}} + \hat{Q}_i^n(0) - \hat{Q}_{-i}^n(0) - \hat{\xi}_i^n + \hat{\xi}_{-i}^n \right) \right]^+. \end{aligned}$$

Let

$$\hat{M}_i^n(t) = \hat{M}_{i,1}^n(t) + \hat{M}_{i,2}^n(t) - \frac{n}{\lambda_{-i}^n} \hat{M}_{i,3}^n(\bar{N}_i^n(t)) + \frac{n}{\lambda_{-i}^n} \hat{M}_{-i,3}^n(\bar{N}_{-i}^n(t)) + \frac{n}{\lambda_{-i}^n} \left( \frac{1}{\sqrt{n}} - \hat{\xi}_i^n + \hat{\xi}_{-i}^n \right).$$

From Lemmas 5.5, 5.7, and 5.8, we have

$$(\hat{M}_1^n, \hat{M}_{-1}^n) \Rightarrow (B_1, -B_1), \quad (5.32)$$

where  $B_1$  is a Brownian motion with drift  $\frac{c}{\lambda}$  and variance  $\lambda(\sigma_1^2 + \sigma_{-1}^2)$ . Next Lemma 5.9 establishes the  $C$ -tightness of  $(\hat{W}_1^n, \hat{W}_{-1}^n)$ . Let  $(\tilde{W}_1, \tilde{W}_{-1})$  be a weak limit of  $(\hat{W}_1^n, \hat{W}_{-1}^n)$  along with a subsequence  $\{n_l\}_{l \in \mathbb{N}}$ . Using Skorohod representation theorem, we can assume  $(\hat{W}_1^{n_l}, \hat{W}_{-1}^{n_l}, \hat{M}_1^{n_l}, \hat{M}_{-1}^{n_l}, \bar{N}_1^{n_l}, \bar{N}_{-1}^{n_l})$  converges to  $(\tilde{W}_1, \tilde{W}_{-1}, B_1, -B_1, \lambda_1 \iota, \lambda_{-1} \iota)$  almost surely and uniformly in compact sets of  $[0, \infty)$ . Define for  $i = \pm 1$  and  $t \geq 0$ ,

$$\epsilon_i^n(t) = \frac{n}{\lambda_{-i}^n} \int_0^t \sqrt{n} F_i^n(\hat{W}_i^n(u-)/\sqrt{n}) d\bar{N}_i^n(u) - \int_0^t H_i(\hat{W}_i^n(u)) du. \quad (5.33)$$

From Lemma 2.4 in [6], we have

$$\epsilon_i^{n_l} \Rightarrow 0, \quad \text{as } l \rightarrow \infty. \quad (5.34)$$

We can now rewrite the waiting time process as follows.

$$\begin{aligned} \hat{W}_i^n(t) = & \left[ \hat{M}_i^n(t) + \frac{n}{\lambda_{-i}^n} \left( \hat{Q}_i^n(0) - \hat{Q}_{-i}^n(0) \right) - \epsilon_i^n(t) + \epsilon_{-i}^n(t) - \int_0^t H_i(\hat{W}_i^n(u)) du \right. \\ & \left. + \int_0^t H_{-i}(\hat{W}_{-i}^n(u)) du \right]^+. \end{aligned}$$

Using Proposition 5.10 (ii) and (5.32), (5.34), we have

$$(\tilde{W}_1, \tilde{W}_{-1}) \stackrel{d}{=} (\Psi_1, \Psi_{-1})(X), \quad (5.35)$$

where  $X$  is a Brownian motion with drift  $\frac{c}{\lambda}$ , variance  $\lambda(\sigma_1^2 + \sigma_{-1}^2)$ , and initial value  $X(0) = q/\lambda$ . Finally, from the uniqueness of  $(\Psi_1, \Psi_{-1})$  in Proposition 5.10 (i), we have  $(\hat{W}_1^n, \hat{W}_{-1}^n) \Rightarrow (\Psi_1, \Psi_{-1})(X)$ .  $\square$

### 5.3. Weak convergence of $\hat{R}_i^n$ and $\hat{G}_i^n$

The following two lemmas show that both  $\hat{R}_i^n$  and  $\hat{G}_i^n$  converge to

$$\lambda_i \int_0^\cdot H_i(\hat{W}_i(s)) ds,$$

where  $\hat{W}_i = \Psi_i(X)$ , with  $X$  defined in Theorem 5.11.

**Lemma 5.12.** *Assume that  $\hat{Q}^n$  is stochastically bounded. Then under Assumptions 3.1, 3.2, and 3.3,*

$$\hat{R}_i^n(\cdot) - \lambda_i \int_0^\cdot H_i(\hat{W}_i^n(s)) ds \Rightarrow 0.$$

*Proof.* From (5.13), for  $t \geq 0$ ,

$$\begin{aligned} & \left| \hat{R}_i^n(t) - \lambda_i \int_0^t H_i(\hat{W}_i^n(s)) ds \right| \\ & \leq \hat{\xi}_i^n + |\hat{M}_{i,3}^n(t)| + \left| \int_0^t \sqrt{n} F_i^n(\hat{W}_i^n(u-)/\sqrt{n}) d\bar{N}_i^n(u) - \lambda_i \int_0^t H_i(\hat{W}_i^n(s)) ds \right|. \end{aligned}$$

The result follows from Lemmas 5.5, 5.7, and (5.34).  $\square$

**Lemma 5.13.** *Assume that  $\hat{Q}^n$  is stochastically bounded. Then under Assumptions 3.1, 3.2, and 3.3,*

$$(R_1^n - G_1^n, R_{-1}^n - G_{-1}^n) \Rightarrow 0.$$

*Proof.* The proof idea is similar to the proof of Proposition 4.1 in [7]. We first show that

$$w_{i,k}^n + t_{i,k}^n \leq w_{i,l}^n + t_{i,l}^n, \quad (5.36)$$

when  $1 \leq k \leq l$ , or  $l \leq k \leq 0$ , or  $k = -Q^n(0) + 1, l = 1$ . Recall that  $t_{i,k}^n = 0$  when  $k \leq 0$ . From (2.4), it is clear that  $w_{1,k}^n \leq w_{1,l}^n$  for  $l \leq k \leq 0$ , and so (5.36) holds for  $l \leq k \leq 0$ . For  $k = -Q^n(0) + 1$  and  $l = 1$ , we have

$$\begin{aligned} w_{1,-Q^n(0)+1}^n &= t_{-1,Q_1^n(0)-R_{1,-Q^n(0)+1}(0)}^n, \\ w_{1,1}^n &= \left( t_{-1,1+Q_1^n(0)-R_1^n(t_{1,1}^n-)+R_{-1}^n(t_{1,1}^n-)}^n - t_{1,1}^n \right)^+. \end{aligned}$$

We note that  $R_{1,-Q^n(0)+1}^n + 1 \geq R_1^n(t_{1,1}^n-)$ , and so

$$t_{-1,Q_1^n(0)-R_{1,-Q^n(0)+1}(0)}^n \leq t_{-1,1+Q_1^n(0)-R_1^n(t_{1,1}^n-)+R_{-1}^n(t_{1,1}^n-)}^n.$$

Thus it is clear that (5.36) holds for  $k = -Q^n(0) + 1$  and  $l = 1$ . Consider now  $k \in \mathbb{N}$ . If  $w_{i,k}^n = 0$ , then (5.36) holds clearly. Suppose now that  $w_{i,k}^n > 0$ . Then

$$w_{i,k}^n + t_{i,k}^n = t_{-i,k+Q_i^n(0)-R_i^n(t_{i,k}^n-)-Q_{-i}^n(0)+R_{-i}^n(t_{i,k}^n-)}^n.$$

Noting that  $R_i^n(t_{i,l}^n-) - R_i^n(t_{i,k}^n-) \leq l - k$ , we have

$$t_{-i,k+Q_i^n(0)-R_i^n(t_{i,k}^n-)-Q_{-i}^n(0)+R_{-i}^n(t_{i,k}^n-)}^n \leq t_{-i,l+Q_i^n(0)-R_i^n(t_{i,l}^n-)-Q_{-i}^n(0)+R_{-i}^n(t_{i,l}^n-)}^n.$$

If  $w_{i,l}^n > 0$ , then

$$w_{i,l}^n + t_{i,l}^n = t_{-i,l+Q_i^n(0)-R_i^n(t_{i,l}^n-)-Q_{-i}^n(0)+R_{-i}^n(t_{i,l}^n-)}^n,$$

and if  $w_{i,l}^n = 0$ , then from Lemma 5.1,

$$w_{i,l}^n + t_{i,l}^n = t_{i,l}^n \geq t_{-i,l+Q_i^n(0)-R_i^n(t_{i,l}^n-)-Q_{-i}^n(0)+R_{-i}^n(t_{i,l}^n-)}^n.$$

Thus (5.36) holds for  $1 \leq k \leq l$ . For  $t \geq 0$ , define

$$\tau_i^n(t) = \inf\{s \geq 0 : s + W_i^n(s) > t\}.$$

Observing that  $w_{i,k}^n + t_{i,k}^n = W_i^n(t_{i,k}^n -) + t_{i,k}^n \leq t$  for all  $t_{i,k}^n < \tau_i^n(t)$ , each customer arriving before time  $\tau_i^n(t)$  should have left the system by time  $t$ . This says  $R_i^n((\tau_i^n(t) - 1/n)^+) \leq G^n(t)$  for  $t \geq 0$ . So we have

$$R_i^n((\tau_i^n(t) - 1/n)^+) \leq G_i^n(t) \leq R_i^n(t), \quad t \geq 0.$$

Next for  $t \geq 0$ , we note that  $\tau_i^n(t) \leq t$ . We then have for  $T > 0$ ,

$$\begin{aligned} & \sup_{0 \leq t \leq T} \left| \hat{R}_i^n(t) - \hat{R}_i^n((\tau_i^n(t) - 1/n)^+) \right| \\ & \leq \sup_{0 \leq t \leq T} \left| \hat{R}_i^n(\tau_i^n(t) + W_i^n(\tau_i^n(t))) - \hat{R}_i^n((\tau_i^n(t) - 1/n)^+) \right| \\ & \leq \sup_{0 \leq t \leq T} \left| \hat{R}_i^n(\tau_i^n(t) + W_i^n(\tau_i^n(t))) - \lambda_i \int_0^{\tau_i^n(t) + W_i^n(\tau_i^n(t))} H_i(\hat{W}_i^n(s)) ds \right| \\ & \quad + \sup_{0 \leq t \leq T} \left| \hat{R}_i^n((\tau_i^n(t) - 1/n)^+) - \lambda_i \int_0^{(\tau_i^n(t) - 1/n)^+} H_i(\hat{W}_i^n(s)) ds \right| \end{aligned} \quad (5.37)$$

$$\begin{aligned} & + \lambda_i \sup_{0 \leq t \leq T} \int_{(\tau_i^n(t) - 1/n)^+}^{\tau_i^n(t) + W_i^n(\tau_i^n(t))} H_i(\hat{W}_i^n(s)) ds \\ & \leq 2 \sup_{0 \leq s \leq T + \sup_{0 \leq u \leq T} W_i^n(u)} \left| \hat{R}_i^n(s) - \lambda_i \int_0^s H_i(\hat{W}_i^n(v)) dv \right| \end{aligned} \quad (5.38)$$

$$+ \lambda_i \sup_{0 \leq s \leq T + \sup_{0 \leq u \leq T} W_i^n(u)} H_i(\hat{W}_i^n(s)) \cdot \sup_{0 \leq s \leq T} [s + W_i^n(s) - (s - 1/n)^+]. \quad (5.39)$$

From Lemmas 5.12 and 5.4, the term in (5.38) converges to 0, and Lemmas 5.4 and 5.9 yields that the term in (5.39) converges to 0. The lemma follows.  $\square$

*Proof of Theorem 4.1.* Recall that  $Q^n(0) \geq 0$ . We then note that for  $t \geq 0$ , from (5.8) and (5.13),

$$\begin{aligned} \hat{W}_1^n(t) &= \left[ \frac{n}{\lambda_{-1}^n} \hat{Q}^{n,+}(0) + \frac{\sqrt{n}}{\lambda_{-1}^n} + \hat{M}_{1,1}^n + \hat{M}_{1,2}^n - \frac{n}{\lambda_{-1}^n} \hat{R}_1^n + \frac{n}{\lambda_{-1}^n} \hat{R}_{-1}^n \right]^+, \\ \hat{W}_{-1}^n(t) &= \left[ -\frac{n}{\lambda_1^n} \hat{Q}^{n,+}(0) + \frac{\sqrt{n}}{\lambda_1^n} + \hat{M}_{-1,1}^n + \hat{M}_{-1,2}^n - \frac{n}{\lambda_1^n} \hat{R}_{-1}^n + \frac{n}{\lambda_1^n} \hat{R}_1^n \right]^+, \end{aligned}$$

and from (2.2),

$$\begin{aligned} \frac{\hat{Q}^{n,+}(t)}{\lambda} &= \left[ \frac{\hat{Q}^n(0)}{\lambda} + \frac{\hat{N}_1^n(t)}{\lambda} - \frac{\hat{N}_{-1}^n(t)}{\lambda} + \frac{(\lambda_1^n - \lambda_{-1}^n)t}{\lambda\sqrt{n}} - \frac{\hat{G}_1^n(t)}{\lambda} + \frac{\hat{G}_{-1}^n(t)}{\lambda} \right]^+, \\ \frac{\hat{Q}^{n,-}(t)}{\lambda} &= \left[ \frac{\hat{Q}^n(0)}{\lambda} + \frac{\hat{N}_1^n(t)}{\lambda} - \frac{\hat{N}_{-1}^n(t)}{\lambda} + \frac{(\lambda_1^n - \lambda_{-1}^n)t}{\lambda\sqrt{n}} - \frac{\hat{G}_1^n(t)}{\lambda} + \frac{\hat{G}_{-1}^n(t)}{\lambda} \right]^-. \end{aligned}$$

Recalling from Lemma 5.5, we have

$$\begin{aligned} \left( \hat{M}_{1,1}^n, \hat{M}_{1,2}^n, \hat{M}_{-1,1}^n, \hat{M}_{-1,2}^n, \frac{\hat{N}_1^n}{\lambda}, \frac{\hat{N}_{-1}^n}{\lambda} \right) &\Rightarrow \left( 0, \frac{c}{\lambda}, 0, -\frac{c}{\lambda}, 0, 0 \right) \iota \\ &\quad + (Y_{-1}, -Y_1, Y_1, -Y_{-1}, -Y_1, -Y_{-1}), \end{aligned} \quad (5.40)$$

where  $\iota : [0, \infty) \rightarrow [0, \infty)$  is the identity map, and  $Y_1(t) = \sigma_1 B_1(\lambda t)$  and  $Y_{-1}(t) = \sigma_{-1} B_{-1}(\lambda t)$  with  $B_1$  and  $B_{-1}$  being two independent standard Brownian motions. Then for  $t \geq 0$ ,

$$\begin{aligned} & \left| \hat{W}_1^n(t) - \frac{\hat{Q}^{n,+}(t)}{\lambda} \right| + \left| \hat{W}_{-1}^n(t) - \frac{\hat{Q}^{n,-}(t)}{\lambda} \right| \\ & \leq \left| \left( \frac{n}{\lambda_{-1}^n} - \frac{1}{\lambda} \right) \hat{Q}^n(0) + \frac{\sqrt{n}}{\lambda_{-1}^n} + \left( \hat{M}_{1,1}^n + \frac{\hat{N}_{-1}^n(t)}{\lambda} \right) + \left( \hat{M}_{1,2}^n - \frac{\hat{N}_1^n(t)}{\lambda} - \frac{(\lambda_1^n - \lambda_{-1}^n)t}{\lambda\sqrt{n}} \right) \right. \\ & \quad \left. + \left( -\frac{n}{\lambda_{-1}^n} \hat{R}_1^n + \frac{\hat{G}_1^n(t)}{\lambda} \right) + \left( \frac{n}{\lambda_{-1}^n} \hat{R}_{-1}^n - \frac{\hat{G}_{-1}^n(t)}{\lambda} \right) \right| \\ & \quad + \left| \left( \frac{n}{\lambda_1^n} - \frac{1}{\lambda} \right) \hat{Q}^n(0) + \frac{\sqrt{n}}{\lambda_1^n} + \left( \hat{M}_{-1,1}^n + \frac{\hat{N}_1^n(t)}{\lambda} \right) + \left( \hat{M}_{-1,2}^n - \frac{\hat{N}_{-1}^n(t)}{\lambda} + \frac{(\lambda_1^n - \lambda_{-1}^n)t}{\lambda\sqrt{n}} \right) \right. \\ & \quad \left. + \left( -\frac{n}{\lambda_1^n} \hat{R}_{-1}^n + \frac{\hat{G}_{-1}^n(t)}{\lambda} \right) + \left( \frac{n}{\lambda_1^n} \hat{R}_1^n - \frac{\hat{G}_1^n(t)}{\lambda} \right) \right|. \end{aligned} \quad (5.41)$$

Combining (5.40) and Lemma 5.13, we have (5.41) converges weakly to 0.  $\square$

*Proof of Theorem 4.2.* From Theorem 4.1,

$$\hat{Q}^n - \lambda(\hat{W}_1^n - \hat{W}_{-1}^n) \Rightarrow 0.$$

Let  $(W_1, W_{-1})$  denote the weak limit of  $(\hat{W}_1^n, \hat{W}_{-1}^n)$  and  $X$  be defined in Theorem 5.11. Then from Theorem 5.11, we see that

$$\hat{Q}^n \Rightarrow \lambda(\Psi_1(X) - \Psi_{-1}(X)) = \lambda X(t) - \lambda \int_0^t H_1(W_1(s)) ds + \lambda \int_0^t H_{-1}(W_{-1}(s)) ds, \quad (5.42)$$

Denote by  $Q$  the weak limit of  $\hat{Q}^n$  in (5.42). Then from Theorem 4.1, we have

$$Q^+ = \lambda W_1, \text{ and } Q^- = \lambda W_{-1}.$$

The result follows.  $\square$

*Proof of Theorem 4.3.* We first obtain the generator of  $Q$  as follows. For  $x \in \mathbb{R}$  and  $f \in C_0^2(\mathbb{R})$ ,

$$\mathcal{A}f(x) = \frac{\lambda^3(\sigma_1^2 + \sigma_{-1}^2)f''(x)}{2} + \left[ c - \lambda H_1\left(\frac{x}{\lambda}\right) 1_{\{x \geq 0\}} + \lambda H_{-1}\left(\frac{x}{\lambda}\right) 1_{\{x < 0\}} \right] f'(x).$$

From Proposition 9.2 in [10], it suffices to verify that  $\int_{\mathbb{R}} \mathcal{A}f(x) \pi(dx) = 0$  for all  $f \in C_0^2(\mathbb{R})$ . Indeed, using integration by parts, we have

$$\begin{aligned} & \int_{[0, \infty)} \mathcal{A}f(x) \pi(dx) \\ & = C_0 \int_{[0, \infty)} \frac{\lambda^3(\sigma_1^2 + \sigma_{-1}^2)f''(x)}{2} \exp \left\{ -\frac{2}{\lambda^3(\sigma_1^2 + \sigma_{-1}^2)} \left( -cx + \lambda^2 \int_0^{\frac{x}{\lambda}} H_1(u) du \right) \right\} dx \\ & \quad + C_0 \int_{[0, \infty)} \left[ c - \lambda H_1\left(\frac{x}{\lambda}\right) \right] f'(x) \exp \left\{ -\frac{2}{\lambda^3(\sigma_1^2 + \sigma_{-1}^2)} \left( -cx + \lambda^2 \int_0^{\frac{x}{\lambda}} H_1(u) du \right) \right\} dx \\ & = -\frac{\lambda^3(\sigma_1^2 + \sigma_{-1}^2)f'(0)}{2} C_0, \end{aligned}$$

and

$$\begin{aligned}
& \int_{(-\infty, 0)} \mathcal{A}f(x)\pi(dx) \\
&= C_0 \int_{(-\infty, 0)} \frac{\lambda^3(\sigma_1^2 + \sigma_{-1}^2)f''(x)}{2} \exp \left\{ -\frac{2}{\lambda^3(\sigma_1^2 + \sigma_{-1}^2)} \left( -cx + \lambda^2 \int_0^{-\frac{x}{\lambda}} H_{-1}(u)du \right) \right\} dx \\
&\quad + C_0 \int_{(-\infty, 0)} \left[ c + \lambda H_{-1} \left( \frac{x}{\lambda} \right) \right] f'(x) \exp \left\{ -\frac{2}{\lambda^3(\sigma_1^2 + \sigma_{-1}^2)} \left( -cx + \lambda^2 \int_0^{-\frac{x}{\lambda}} H_{-1}(u)du \right) \right\} dx \\
&= \frac{\lambda^3(\sigma_1^2 + \sigma_{-1}^2)f'(0)}{2} C_0.
\end{aligned}$$

Thus  $\pi$  is a stationary distribution of  $Q$ . Finally, the uniqueness of  $\pi$  follows from the irreducibility of  $Q$ .  $\square$

#### 5.4. Proof of Corollary 4.4

The goal is to prove  $\hat{Q}^n$  is  $C$ -tight under Assumption 3.1, 3.2, and 3.4. We first follow [15] to construct the compensator for the abandonment process  $G_i^n$ . Define the potential waiting time process for the customers of Class  $i$  as follows: For  $k \in \mathbb{N}$ ,

$$\tilde{w}_{i,k}^n(t) = \begin{cases} \max\{t - t_{i,k}^n, 0\}, & \text{if } t - t_{i,k}^n < d_{i,k}^n, \\ d_{i,k}^n, & \text{otherwise.} \end{cases}$$

We note that for  $k \in \mathbb{N}$ ,  $\tilde{w}_{i,k}^n(t)$  represents the amount of time spent by the  $k^{th}$  customer of Class  $i$  since entering the system, and remains constant at  $d_{i,k}^n$  once the time spent reaches the patience time, and for  $k \in -\mathbb{N} \cup \{0\}$ ,  $\tilde{w}_{i,k}^n$  represents the potential waiting time process of the  $(-k+1)^{st}$  customer who enters the system before time 0 (if such customer exists). The following measure, which is called the potential queue measure, assigns a unit mass to the potential waiting time of each customer of Class  $i$  that has entered the system by time  $t$  and whose potential waiting time has not yet reached the patience time.

$$\eta_{i,t}^n(dx) = \sum_{k=-Q_i^n(0)+1}^{N_i^n(t)} \delta_{\tilde{w}_{i,k}^n}(dx) 1_{\{\tilde{w}_{i,k}^n(t) < d_{i,k}^n\}} = \sum_{k=-Q_i^n(0)+1}^{N_i^n(t)} \delta_{\tilde{w}_{i,k}^n}(dx) 1_{\left\{ \frac{d\tilde{w}_{i,k}^n(t+)}{dt} > 0 \right\}}.$$

The number of Class  $i$  customers at time  $t$  can then be formulated as follows:

$$Q_i^n(t) = \sum_{k=-Q_i^n(0)+1}^{N_i^n(t)} 1_{\{\tilde{w}_{i,k}^n(t) \leq \chi_i^n(t), \tilde{w}_{i,k}^n(t) < d_{i,k}^n\}} = \eta_{i,t}^n[0, \chi_i^n(t)],$$

where  $\chi_i^n(t) = \inf\{x > 0 : \eta_{i,t}^n[0, x] \geq Q_i^n(t)\}$  which is the waiting time of the head-of-the-line Class  $i$  customer in the queue at time  $t$ . The abandonment process of Class  $i$  at time  $t$  becomes

$$G_i^n(t) = \sum_{k=-Q_i^n(0)+1}^{N_i^n(t)} \sum_{s \in [0, t]} 1_{\left\{ \tilde{w}_{i,k}^n(s) \leq \chi_i^n(s-), \frac{d\tilde{w}_{i,k}^n(s-)}{dt} > 0, \frac{d\tilde{w}_{i,k}^n(s+)}{dt} = 0 \right\}}.$$

Recall from Assumption 3.4 that for  $i = \pm 1$ , the patience time distribution function  $F_i^n$  has density function  $f_i^n$  and hazard rate function  $h_i^n = f_i^n / (1 - F_i^n)$  on  $[0, M_i^n)$ , where  $M_i^n = \sup\{s \geq 0 : F_i^n(s) < 1\}$ . Furthermore, for  $t \in [0, M_i^n)$ ,  $\sup_{n \in \mathbb{N}} \|h_i^n\|_t < \infty$ , and  $h_i^n(t/\sqrt{n}) \rightarrow h_i(t)$ , where  $h_i$  is a nonnegative measurable function. Define for  $t \geq 0$ ,

$$A_i^n(t) = \int_0^t \left( \int_{[0, M_i^n)} 1_{[0, \chi_i^n(s-)]}(x) h_i^n(x) \eta_{i,s}^n(dx) \right) ds,$$

and

$$\begin{aligned} \mathcal{G}_{i,t}^n = \sigma \{ & Q^n(0); (N_i^n(s), N_{-i}^n(s)), 0 \leq s \leq t; \\ & (w_{i,l}^n, d_{i,l}^n), l = -Q_i^n(0) + 1, \dots, 0, 1, \dots, N_i(t); \\ & (w_{-i,l}^n, d_{-i,l}^n), l = -Q_{-i}^n(0) + 1, \dots, 0, 1, \dots, N_{-i}^n(t); \\ & t_{-i,l}^n, l = 1, 2, \dots, N_i^n(t) + Q_i^n(0) - R_i^n(t) - Q_{-i}^n(0) + R_{-i}^n(t) \}. \end{aligned}$$

**Lemma 5.14.** *For  $n \in \mathbb{N}$ , the process  $A_i^n$  is the  $\{\mathcal{G}_{i,t}^n\}$  compensator of the process  $G_i^n$ , and in particular,  $G_i^n - A_i^n$  is a  $\{\mathcal{G}_{i,t}^n\}$  local martingale.*

*Proof.* First it can be seen that  $(Q^n, G_i^n, A_i^n) \in \{\mathcal{G}_{i,t}^n\}$ . Then the rest of the proof is essentially the same as that of Lemma 5.4 in [15].  $\square$

**Lemma 5.15.** *For  $t \geq 0$ ,*

$$A_i^n(t) = \sum_{k=-Q_i^n(0)+1}^{N_i^n(t)} \int_0^{(t-t_{i,k}^n) \wedge w_{i,k}^n \wedge d_{i,k}^n} h_i^n(u) du.$$

*Proof.* The following proof is adapted from [25]. We first observe that for  $t \geq 0$ ,

$$1_{[0, \chi_i^n(s-)]}(x) \eta_{i,s}^n(dx) = \sum_{k=-Q_i^n(0)+1}^{N_i^n(s)} 1_{\{s-t_{i,k}^n < w_{i,k}^n\}}(x) 1_{\{s-t_{i,k}^n < d_{i,k}^n\}}(x) \delta_{s-t_{i,k}^n}(dx),$$

and so

$$\int_{[0, M_i^n)} 1_{[0, \chi_i^n(s-)]}(x) h_i^n(x) \eta_{i,s}^n(dx) = \sum_{k=-Q_i^n(0)+1}^{N_i^n(s)} h_i^n(s - t_{i,k}^n) 1_{\{s-t_{i,k}^n < w_{i,k}^n \wedge d_{i,k}^n\}}.$$

Finally, we have for  $t \geq 0$ ,

$$\begin{aligned} A_i^n(t) &= \int_0^t \sum_{k=-Q_i^n(0)+1}^{N_i^n(s)} h_i^n(s - t_{i,k}^n) 1_{\{s-t_{i,k}^n < w_{i,k}^n \wedge d_{i,k}^n\}} ds \\ &= \sum_{k=-Q_i^n(0)+1}^{N_i^n(t)} \int_{t_{i,k}^n}^t h_i^n(s - t_{i,k}^n) 1_{\{s-t_{i,k}^n < w_{i,k}^n \wedge d_{i,k}^n\}} ds \\ &= \sum_{k=-Q_i^n(0)+1}^{N_i^n(t)} \int_0^{(t-t_{i,k}^n) \wedge w_{i,k}^n \wedge d_{i,k}^n} h_i^n(u) du. \end{aligned}$$

$\square$



**Lemma 5.16.** Assume  $\hat{Q}^n(0)$  converges weakly to some random variable  $q$ . Then  $\hat{Q}^n$  is  $C$ -tight.

*Proof.* For  $t \geq 0$ , we note that

$$\hat{Q}^n(t) = \hat{Q}^n(0) + \hat{N}_1^n(t) - \hat{N}_{-1}^n(t) + \frac{(\lambda_1^n - \lambda_{-1}^n)t}{\sqrt{n}} - \hat{G}_1^n(t) + \hat{G}_{-1}^n(t).$$

Define for  $t \geq 0$ ,

$$\hat{A}_i^n(t) = \frac{1}{\sqrt{n}} A_i^n(t), \text{ and } \hat{M}_i^n(t) = \hat{G}_i^n(t) - \hat{A}_i^n(t).$$

From Lemmas 5.14 and 5.15, the quadratic variation of the local martingale  $\hat{M}_i^n$  is

$$\begin{aligned} [\hat{M}_i^n]_t &= \frac{1}{n} A_i^n(t) = \frac{1}{n} \int_0^t \sum_{k=-Q_i^n(0)+1}^{N_i^n(s)} h_i^n(s - t_{i,k}^n) 1_{\{s-t_{i,k}^n < w_{i,k}^n \wedge d_{i,k}^n\}} ds \\ &\leq \frac{\|h_i^n\|_t}{n} \int_0^t \sum_{k=-Q_i^n(0)+1}^{N_i^n(s)} 1_{\{s-t_{i,k}^n < w_{i,k}^n \wedge d_{i,k}^n\}} ds \\ &= \frac{\|h_i^n\|_t}{n} \int_0^t Q_i^n(s) ds = \|h_i^n\|_t \int_0^t \bar{Q}_i^n(s) ds \rightarrow 0, \text{ in probability.} \end{aligned} \tag{5.43}$$

Thus from Lemma 5.6, we have  $\hat{M}_i^n \Rightarrow 0$ . Furthermore, we have for  $t \geq 0$ ,

$$\|\hat{Q}^n\|_t \leq |\hat{Q}^n(0)| + \sum_{i=\pm 1} (\|\hat{N}_i^n\|_t + \|\hat{M}_i^n\|_t) + \frac{|\lambda_1^n - \lambda_{-1}^n|t}{\sqrt{n}} + \sum_{i=\pm 1} \|h_i^n\|_t \int_0^t \|\hat{Q}^n\|_s ds.$$

Using Gronwall's inequality, we have

$$\|\hat{Q}^n\|_t \leq e^{\sum_{i=\pm 1} \|h_i^n\|_t} \left( |\hat{Q}^n(0)| + \sum_{i=\pm 1} (\|\hat{N}_i^n\|_t + \|\hat{M}_i^n\|_t) + \frac{|\lambda_1^n - \lambda_{-1}^n|t}{\sqrt{n}} \right). \tag{5.44}$$

Next using similar argument as in (5.43), we have for  $0 \leq s \leq t < \infty$ ,

$$\begin{aligned} \hat{A}_i^n(t) - \hat{A}_i^n(s) &= \frac{1}{\sqrt{n}} \int_s^t \sum_{k=-Q_i^n(0)+1}^{N_i^n(u)} h_i^n(u - t_{i,k}^n) 1_{\{u-t_{i,k}^n < w_{i,k}^n \wedge d_{i,k}^n\}} du \\ &\leq \|h_i^n\|_{s,t} \int_s^t \hat{Q}_i^n(u) du, \end{aligned}$$

and therefore,

$$\begin{aligned} &|\hat{Q}^n(t) - \hat{Q}^n(s)| \\ &\leq \sum_{i=\pm 1} (|\hat{N}_i^n(t) - \hat{N}_i^n(s)| + |\hat{M}_i^n(t) - \hat{M}_i^n(s)| + |\hat{A}_i^n(t) - \hat{A}_i^n(s)|) + \frac{(\lambda_1^n - \lambda_{-1}^n)(t-s)}{\sqrt{n}} \\ &\leq \sum_{i=\pm 1} (|\hat{N}_i^n(t) - \hat{N}_i^n(s)| + |\hat{M}_i^n(t) - \hat{M}_i^n(s)|) + \sum_{i=\pm 1} \|h_i^n\|_{s,t} \int_s^t |\hat{Q}^n(u)| du + \frac{(\lambda_1^n - \lambda_{-1}^n)(t-s)}{\sqrt{n}} \\ &\leq \sum_{i=\pm 1} (|\hat{N}_i^n(t) - \hat{N}_i^n(s)| + |\hat{M}_i^n(t) - \hat{M}_i^n(s)|) + (t-s) \|\hat{Q}^n\|_t \sum_{i=\pm 1} \|h_i^n\|_{s,t} + \frac{(\lambda_1^n - \lambda_{-1}^n)(t-s)}{\sqrt{n}}. \end{aligned}$$

The above estimate and (5.44) implies that  $\hat{Q}^n$  is  $C$ -tight on noting that  $\hat{N}_i^n$  and  $\hat{M}_i^n$  are weakly convergent, and that the jump seize of  $\hat{Q}_i^n$  is bounded by  $\frac{1}{\sqrt{n}}$ .  $\square$

*Proof of Corollary 4.4.* This follows immediately from Lemma 5.16, Theorem 4.2, and Remark 3.1 (iv).  $\square$

## Appendix

*Proof of Proposition 5.10.* Fix  $T > 0$ , and let  $\kappa > 0$  be such that for  $t_1, t_2 \in [0, T]$ ,

$$|H_1(t_1) - H_1(t_2)| \leq \kappa|t_1 - t_2|, \text{ and } |H_{-1}(t_1) - H_{-1}(t_2)| \leq \kappa|t_1 - t_2|.$$

(i) Let  $w(t) = w_1(t) + w_{-1}(t)$ ,  $t \in [0, T]$ . We first show that there exists  $M \in (0, \infty)$  such that  $\|w\|_T \leq M$ . We first note that for  $t \geq 0$ ,

$$\begin{aligned} w(t) &= \left| x(t) - \int_0^t H_1(w_1(s))ds + \int_0^t H_{-1}(w_{-1}(s))ds \right| \\ &\leq |x(t)| + \int_0^t (H_1 + H_{-1})(w(s))ds. \end{aligned}$$

Let  $H(x) = H_1(x) + H_{-1}(x) + 1$ ,  $x \in [0, \infty)$ . Then we have for  $t \in [0, T]$ ,

$$\|w\|_t \leq \|x\|_t + \int_0^t H(\|w\|_s)ds. \quad (5.45)$$

Let  $y(t) = \int_0^t H(\|w\|_s)ds$ ,  $t \in [0, T]$ . Then from (5.45), for  $t \in [0, T]$ ,

$$y'(t) = H(\|w\|_t) \leq H(\|x\|_T + y(t)),$$

Integrating the above equation on  $[0, T]$ , we have

$$\int_0^T \frac{y'(t)}{H(\|x\|_T + y(t))}dt \leq T, \quad (5.46)$$

and so

$$\int_{\|x\|_T}^{\|x\|_T + y(T)} \frac{du}{H(u)} \leq T. \quad (5.47)$$

Now define for  $t \in [0, T]$ ,

$$\Phi(t) = \int_0^t \frac{du}{H(u)}.$$

Since  $H$  is strictly positive, nondecreasing, and Lipschitz continuous on  $[0, T]$ , it is clear that  $\Phi$  is increasing and differentiable. Finally, we have from (5.47),

$$\Phi(\|x\|_T + y(T)) - \Phi(\|x\|_T) \leq T,$$

and so

$$y(T) \leq \Phi^{-1}(\Phi(\|x\|_T) + T) - \|x\|_T.$$

Now from (5.45),

$$\|w\|_T \leq \Phi^{-1}(\Phi(\|x\|_T) + T). \quad (5.48)$$

Denoting the RHS of (5.48) by  $M$ , we have  $\|w\|_T \leq M$ . We next use Picard iteration to show the existence of  $(w_1, w_{-1})$ . Define a sequence of  $(w_1^n, w_{-1}^n)_{n=0}^\infty$  recursively as follows. Let  $(w_1^0, w_{-1}^0) = 0$ , and for  $n \geq 1$  and  $t \in [0, T]$ ,

$$\begin{aligned} w_1^n(t) &= \left[ x(t) - \int_0^t H_1(w_1^{n-1}(s))ds + \int_0^t H_{-1}(w_{-1}^{n-1}(s))ds \right]^+, \\ w_{-1}^n(t) &= \left[ x(t) - \int_0^t H_1(w_1^{n-1}(s))ds + \int_0^t H_{-1}(w_{-1}^{n-1}(s))ds \right]^-. \end{aligned}$$

Then for  $s \in [0, T]$ ,

$$\|w_1^1\|_s \leq \|x\|_s, \quad \|w_{-1}^1\|_s \leq \|x\|_s, \quad (5.49)$$

for  $n \geq 1$ ,

$$\begin{aligned} &\|w_1^n - w_1^{n-1}\|_s \\ &= \sup_{0 \leq t \leq s} \left| \left[ x(t) - \int_0^t H_1(w_1^{n-1}(s))ds + \int_0^t H_{-1}(w_{-1}^{n-1}(s))ds \right]^+ \right. \\ &\quad \left. - \left[ x(t) - \int_0^t H_1(w_1^{n-2}(s))ds + \int_0^t H_{-1}(w_{-1}^{n-2}(s))ds \right]^+ \right| \\ &\leq \sup_{0 \leq t \leq s} \left| \int_0^t -H_1(w_1^{n-1}(s)) + H_1(w_1^{n-2}(s))ds + \int_0^t H_{-1}(w_{-1}^{n-1}(s)) - H_{-1}(w_{-1}^{n-2}(s))ds \right| \\ &\leq \int_0^s |H_1(w_1^{n-1}(s)) - H_1(w_1^{n-2}(s))|ds + \int_0^s |H_{-1}(w_{-1}^{n-1}(s)) - H_{-1}(w_{-1}^{n-2}(s))|ds \quad (5.50) \\ &\leq \kappa s (\|w_1^{n-1} - w_1^{n-2}\|_s + \|w_{-1}^{n-1} - w_{-1}^{n-2}\|_s), \end{aligned}$$

and using similar analysis, we have for  $n \geq 1$  and  $s \in [0, T]$ ,

$$\|w_{-1}^n - w_{-1}^{n-1}\|_s \leq \kappa s (\|w_1^{n-1} - w_1^{n-2}\|_s + \|w_{-1}^{n-1} - w_{-1}^{n-2}\|_s).$$

Thus for  $s \in [0, T]$ ,

$$\|w_1^1\|_s + \|w_{-1}^1\|_s \leq 2\|x\|_s, \quad (5.51)$$

and

$$\|w_1^n - w_1^{n-1}\|_s + \|w_{-1}^n - w_{-1}^{n-1}\|_s \leq 2\kappa s (\|w_1^{n-1} - w_1^{n-2}\|_s + \|w_{-1}^{n-1} - w_{-1}^{n-2}\|_s), \quad n \geq 2. \quad (5.52)$$

The rest of the proof is similar to the proof of Lemma 1 in [26]. Choose  $\delta > 0$  such that  $2\kappa\delta < 1$ , and then partition  $[0, T]$  into  $\lfloor T/\delta \rfloor + 1$  subintervals  $[y_j, y_{j+1}]$ ,  $j = 0, 1, \dots, \lfloor T/\delta \rfloor$ , where  $y_0 = 0$ ,  $y_j = j\delta$ ,  $j = 1, \dots, \lfloor T/\delta \rfloor$ , and  $y_{\lfloor T/\delta \rfloor + 1} = T$ . Let  $c = 2\|x\|_T$ . We first observe that for  $n \geq 1$ ,

$$\|w_1^n - w_1^{n-1}\|_\delta + \|w_{-1}^n - w_{-1}^{n-1}\|_\delta \leq 2\kappa\delta (\|w_1^{n-1} - w_1^{n-2}\|_\delta + \|w_{-1}^{n-1} - w_{-1}^{n-2}\|_\delta),$$

and repeating the iterations, we have

$$\|w_1^n - w_1^{n-1}\|_\delta + \|w_{-1}^n - w_{-1}^{n-1}\|_\delta \leq (2\kappa\delta)^{n-1} \cdot 2\|x\|_\delta \leq (2\kappa\delta)^{n-1}c. \quad (5.53)$$

In the following, we use induction to show that for  $j = 1, \dots, \lfloor T/\delta \rfloor + 1$ ,

$$\|w_1^n - w_1^{n-1}\|_{y_j} + \|w_{-1}^n - w_{-1}^{n-1}\|_{y_j} \leq j(n-1)^j(2\kappa\delta)^{n-1}c. \quad (5.54)$$

From (5.53), the above inequality (5.54) holds for  $j = 1$ . Suppose (5.54) holds for  $j = 1, \dots, k$ . Now from (5.50),

$$\begin{aligned} & \|w_1^n - w_1^{n-1}\|_{y_{k+1}} + \|w_{-1}^n - w_{-1}^{n-1}\|_{y_{k+1}} \\ & \leq 2 \sum_{j=1}^k \int_{y_{j-1}}^{y_j} |H_1(w_1^{n-1}(s)) - H_1(w_1^{n-2}(s))| + |H_{-1}(w_{-1}^{n-1}(s)) - H_{-1}(w_{-1}^{n-2}(s))| ds \\ & \quad + 2 \int_{y_k}^{y_{k+1}} |H_1(w_1^{n-1}(s)) - H_1(w_1^{n-2}(s))| + |H_{-1}(w_{-1}^{n-1}(s)) - H_{-1}(w_{-1}^{n-2}(s))| ds \\ & \leq 2\kappa\delta \sum_{j=1}^k (\|w_1^{n-1} - w_1^{n-2}\|_{y_j} + \|w_{-1}^{n-1} - w_{-1}^{n-2}\|_{y_j}) + 2\kappa\delta (\|w_1^{n-1} - w_1^{n-2}\|_{y_{k+1}} + \|w_{-1}^{n-1} - w_{-1}^{n-2}\|_{y_{k+1}}) \\ & \leq 2\kappa\delta \sum_{j=1}^k j(n-2)^j(2\kappa\delta)^{n-2}c + 2\kappa\delta (\|w_1^{n-1} - w_1^{n-2}\|_{y_{k+1}} + \|w_{-1}^{n-1} - w_{-1}^{n-2}\|_{y_{k+1}}). \end{aligned}$$

There exists  $n_0 \in \mathbb{N}$  such that when  $n \geq n_0$ ,

$$\sum_{j=1}^k j(n-2)^j \leq k(n-1)^k.$$

Thus when  $n \geq n_0$ ,

$$\begin{aligned} & \|w_1^n - w_1^{n-1}\|_{y_{k+1}} + \|w_{-1}^n - w_{-1}^{n-1}\|_{y_{k+1}} \\ & \leq ck(n-1)^k(2\kappa\delta)^{n-1} + 2\kappa\delta (\|w_1^{n-1} - w_1^{n-2}\|_{y_{k+1}} + \|w_{-1}^{n-1} - w_{-1}^{n-2}\|_{y_{k+1}}). \end{aligned} \quad (5.55)$$

From (5.55) and (5.51), we have

$$\begin{aligned} & \|w_1^2 - w_1^1\|_{y_{k+1}} + \|w_{-1}^2 - w_{-1}^1\|_{y_{k+1}} \\ & \leq ck \cdot 2\kappa\delta + 2\kappa\delta (\|w_1^1\|_{y_{k+1}} + \|w_{-1}^1\|_{y_{k+1}}) \\ & \leq c(k+1) \cdot 2\kappa\delta. \end{aligned}$$

Iterating (5.55) shows that for large enough  $n$ ,

$$\begin{aligned} & \|w_1^n - w_1^{n-1}\|_{y_{k+1}} + \|w_{-1}^n - w_{-1}^{n-1}\|_{y_{k+1}} \\ & \leq c(2\kappa\delta)^{n-1} \left( k \sum_{j=1}^{n-1} j^k + 1 \right) \leq c(2\kappa\delta)^{n-1} (k(n-1)^{k+1} + 1) \\ & \leq c(2\kappa\delta)^{n-1} (k+1)(n-1)^{k+1}. \end{aligned}$$

Thus we show (5.54) holds. When  $j = \lfloor T/\delta \rfloor + 1$ ,

$$\|w_1^n - w_1^{n-1}\|_T + \|w_{-1}^n - w_{-1}^{n-1}\|_T \leq c(2\kappa\delta)^{n-1}(\lfloor T/\delta \rfloor + 2)(n-1)^{\lfloor T/\delta \rfloor + 2} \rightarrow 0, \text{ as } n \rightarrow \infty.$$

So  $(w_1^n, w_{-1}^n)$  converges and denote the limit function by  $(w_1^*, w_{-1}^*) = \lim_{n \rightarrow \infty} (w_1^n, w_{-1}^n)$ . From (5.29) and (5.30), define the functional  $G = (G_1, G_{-1}) : D([0, T], \mathbb{R}_+^2) \rightarrow D([0, T], \mathbb{R}_+^2)$  as follows: For  $t \in [0, T]$ ,

$$\begin{aligned} G(w)_1(t) &= \left[ x(t) - \int_0^t H_1(w_1(s))ds + \int_0^t H_{-1}(w_{-1}(s))ds \right]^+, \\ G(w)_{-1}(t) &= \left[ x(t) - \int_0^t H_1(w_1(s))ds + \int_0^t H_{-1}(w_{-1}(s))ds \right]^-. \end{aligned}$$

Thus  $w = G(w)$ . We next note that  $G$  is Lipschitz continuous, i.e., for  $s \in [0, T]$

$$\|G(w) - G(\tilde{w})\|_s \leq 2\kappa T \|w - \tilde{w}\|_s.$$

Using Banach fixed point theorem, we have  $w = (w_1^*, w_{-1}^*)$ .

(ii) Let  $x^n, x \in D([0, \infty), \mathbb{R})$  such that  $x^n \rightarrow x$  in the Skorohod  $J_1$  topology. Fix  $T \in (0, \infty)$ . Then there exists a sequence of strictly increasing homeomorphism  $\{\lambda^n\}$  on  $[0, \infty)$  such that

$$\sup_{0 \leq t \leq T} |x^n(\lambda^n(t)) - x(t)| \vee \sup_{0 \leq t \leq T} |\lambda^n(t) - t| \rightarrow 0, \text{ as } n \rightarrow \infty. \quad (5.56)$$

Let

$$w^n \equiv (w_1^n, w_{-1}^n) = (\Psi_1(x^n), \Psi_{-1}(x^n)), \text{ and } w \equiv (w_1, w_{-1}) = (\Psi_1(x), \Psi_{-1}(x)).$$

It is clear that, from (5.56), there exists  $M > 0$  such that

$$\sup_{n \in \mathbb{N}} \|x^n\|_T + \|x\|_T \leq M,$$

and from (5.48), there exists  $\tilde{M} > 0$  such that

$$\sup_{n \in \mathbb{N}} \|w^n\|_T + \|w\|_T \leq \tilde{M}.$$

Let  $\kappa > 0$  be the Lipschitz constant for  $H_1$  and  $H_{-1}$  on  $[0, \tilde{M}]$ . For  $u \in [0, T]$ ,

$$\begin{aligned} \sup_{0 \leq t \leq u} |w_1^n(\lambda^n(t)) - w_1(t)| &= \sup_{0 \leq t \leq u} \left| \left[ x^n(\lambda^n(t)) - \int_0^{\lambda^n(t)} H_1(w_1^n(s))ds + \int_0^{\lambda^n(t)} H_{-1}(w_{-1}^n(s))ds \right]^+ \right. \\ &\quad \left. - \left[ x(t) - \int_0^t H_1(w_1(s))ds + \int_0^t H_{-1}(w_{-1}(s))ds \right]^+ \right| \\ &\leq \sup_{0 \leq t \leq u} |x^n(\lambda^n(t)) - x(t)| + \sup_{0 \leq t \leq u} \left| \int_0^{\lambda^n(t)} H_1(w_1^n(s))ds - \int_0^t H_1(w_1(s))ds \right| \\ &\quad + \sup_{0 \leq t \leq u} \left| \int_0^{\lambda^n(t)} H_{-1}(w_{-1}^n(s))ds - \int_0^t H_{-1}(w_{-1}(s))ds \right|. \end{aligned}$$

We next observe that for  $u \in [0, T]$ ,

$$\begin{aligned}
& \int_0^{\lambda^n(t)} H_1(w_1^n(s))ds - \int_0^t H_1(w_1(s))ds \\
&= \int_0^{\lambda^n(t)} H_1(w_1^n(s))ds - \int_0^t H_1(w_1^n(\lambda^n(s)))ds + \int_0^t H_1(w_1^n(\lambda^n(s)))ds - \int_0^t H_1(w_1(s))ds \\
&= \int_0^t H_1(w_1^n(s)) - H_1(w_1^n(\lambda^n(s)))ds + \int_t^{\lambda^n(t)} H_1(w_1^n(s))ds + \int_0^t H_1(w_1^n(\lambda^n(s))) - H_1(w_1(s))ds.
\end{aligned}$$

We note that from the Lipschitz continuity of  $H_1$  and (5.29),

$$\begin{aligned}
& \sup_{0 \leq t \leq T} \left| \int_0^t H_1(w_1^n(s)) - H_1(w_1^n(\lambda^n(s)))ds \right| \\
& \leq T\kappa \sup_{0 \leq t \leq T} |w_1^n(t) - w_1^n(\lambda^n(t))| \\
& \leq T\kappa \sup_{0 \leq t \leq T} \left( |x^n(t) - x^n(\lambda^n(t))| + \left| \int_t^{\lambda^n(t)} H_1(w_1(s))ds \right| + \left| \int_t^{\lambda^n(t)} H_{-1}(w_{-1}(s))ds \right| \right) \\
& \leq \kappa T \sup_{0 \leq t \leq T} |x^n(t) - x^n(\lambda^n(t))| + 2\kappa^2 T \tilde{M} \sup_{0 \leq t \leq T} |\lambda^n(t) - t|
\end{aligned}$$

Thus

$$\begin{aligned}
& \sup_{0 \leq t \leq u} \left| \int_0^{\lambda^n(t)} H_1(w_1^n(s))ds - \int_0^t H_1(w_1(s))ds \right| \\
& \leq \sup_{0 \leq t \leq T} \left| \int_0^t H_1(w_1^n(s)) - H_1(w_1^n(\lambda^n(s)))ds \right| + \sup_{0 \leq t \leq T} \left| \int_t^{\lambda^n(t)} H_1(w_1^n(s))ds \right| \\
& \quad + \sup_{0 \leq t \leq T} \left| \int_0^t H_1(w_1^n(\lambda^n(s))) - H_1(w_1(s))ds \right| \\
& \leq \kappa T \sup_{0 \leq t \leq T} |x^n(t) - x^n(\lambda^n(t))| + 2\kappa^2 T \tilde{M} \sup_{0 \leq t \leq T} |\lambda^n(t) - t| + \kappa \tilde{M} \sup_{0 \leq t \leq T} |\lambda^n(t) - t| \\
& \quad + 2\kappa \int_0^T \sup_{0 \leq s \leq t} |w_1^n(\lambda^n(s)) - w_1(s)|dt \\
& = \kappa T \sup_{0 \leq t \leq T} |x^n(t) - x^n(\lambda^n(t))| + (2\kappa^2 + \kappa)T \tilde{M} \sup_{0 \leq t \leq T} |\lambda^n(t) - t| + 2\kappa \int_0^T \sup_{0 \leq s \leq t} |w_1^n(\lambda^n(s)) - w_1(s)|dt,
\end{aligned}$$

and similarly,

$$\begin{aligned}
& \sup_{0 \leq t \leq u} \left| \int_0^{\lambda^n(t)} H_{-1}(w_{-1}^n(s))ds - \int_0^t H_{-1}(w_{-1}(s))ds \right| \\
& \leq \kappa T \sup_{0 \leq t \leq T} |x^n(t) - x^n(\lambda^n(t))| + (2\kappa^2 + \kappa)T \tilde{M} \sup_{0 \leq t \leq T} |\lambda^n(t) - t| + 2\kappa \int_0^T \sup_{0 \leq s \leq t} |w_{-1}^n(\lambda^n(s)) - w_{-1}(s)|dt.
\end{aligned}$$

Combining the above estimates, we have

$$\begin{aligned}
& \sup_{0 \leq t \leq T} |w_1^n(\lambda^n(t)) - w_1(t)| \leq (1 + 2\kappa T) \sup_{0 \leq t \leq T} |x^n(\lambda^n(t)) - x(t)| \\
& + 2(2\kappa^2 + \kappa)T \tilde{M} \sup_{0 \leq t \leq T} |\lambda^n(t) - t| + 2\kappa \int_0^T \sup_{0 \leq s \leq t} |w_1^n(\lambda^n(s)) - w_1(s)| + \sup_{0 \leq s \leq t} |w_{-1}^n(\lambda^n(s)) - w_{-1}(s)|dt.
\end{aligned}$$

A similar analysis yields that

$$\begin{aligned} \sup_{0 \leq t \leq T} |w_{-1}^n(\lambda^n(t)) - w_{-1}(t)| &\leq (1 + 2\kappa T) \sup_{0 \leq t \leq T} |x^n(\lambda^n(t)) - x(t)| \\ &+ 2(2\kappa^2 + \kappa)T\tilde{M} \sup_{0 \leq t \leq T} |\lambda^n(t) - t| + 2\kappa \int_0^T \sup_{0 \leq s \leq t} |w_1^n(\lambda^n(s)) - w_1(s)| + \sup_{0 \leq s \leq t} |w_{-1}^n(\lambda^n(s)) - w_{-1}(s)| dt. \end{aligned}$$

Finally, we have for  $u \in [0, T]$ ,

$$\begin{aligned} \sup_{0 \leq t \leq u} |w_1^n(\lambda^n(t)) - w_1(t)| + \sup_{0 \leq t \leq u} |w_{-1}^n(\lambda^n(t)) - w_{-1}(t)| &\leq 2(1 + 2\kappa T) \sup_{0 \leq t \leq u} |x^n(\lambda^n(t)) - x(t)| \\ &+ 4(2\kappa^2 + \kappa)T\tilde{M} \sup_{0 \leq t \leq u} |\lambda^n(t) - t| + 4\kappa \int_0^T \sup_{0 \leq s \leq t} |w_1^n(\lambda^n(s)) - w_1(s)| + \sup_{0 \leq s \leq t} |w_{-1}^n(\lambda^n(s)) - w_{-1}(s)| dt. \end{aligned}$$

Using Gronwall's inequality, we have for  $u \in [0, T]$ ,

$$\begin{aligned} &\sup_{0 \leq t \leq T} |w_1^n(\lambda^n(t)) - w_1(t)| + \sup_{0 \leq t \leq u} |w_{-1}^n(\lambda^n(t)) - w_{-1}(t)| \\ &\leq \left( 2(1 + 2\kappa T) \sup_{0 \leq t \leq T} |x^n(\lambda^n(t)) - x(t)| + 4(2\kappa^2 + \kappa)T\tilde{M} \sup_{0 \leq t \leq T} |\lambda^n(t) - t| \right) e^{4\kappa u}. \end{aligned}$$

Finally, from the above estimate, we see that

$$\sup_{0 \leq t \leq T} |w^n(\lambda^n(t)) - w(t)| \vee \sup_{0 \leq t \leq T} |\lambda^n(t) - t| \rightarrow 0, \text{ as } n \rightarrow \infty.$$

□

## References

- [1] P. Afeche, A. Diamant, and J. Milner, *Double-sided batch queues with abandonment: Modeling crossing networks*, Operations Research **62** (2014), no. 5, 1179–1201.
- [2] O.J. Boxma, I. David, D. Perry, and W. Stadje, *A new look at organ transplantation models and double matching queues*, Probability in the Engineering and Informational Sciences **25** (2011), 135–155.
- [3] B. W. Conolly, P. R. Parthasarathy, and N. Selvaraju, *Double-ended queues with impatience*, Computers & Operations Research **29** (2002), no. 14, 2053–2072.
- [4] J. G. Dai and S. He, *Customer abandonment in many-server queues*, Math. Oper. Res. **35** (2010), no. 2, 347–362.
- [5] J. G. Dai, S. He, and T. Tezcan, *Many-server diffusion limits for  $G/Ph/n + GI$  queues*, Ann. Appl. Probab. **20** (2010), no. 5, 1854–1890.
- [6] J. G. Dai and R. J. Williams, *Existence and uniqueness of semimartingale reflecting brownian motions in convex polyhedrons*, Theory of Probability & Its Applications **40** (1996), no. 1, 1–40.
- [7] Jim Dai and Shuangchi He, *Customer abandonment in many-server queues*, Mathematics of Operations Research **35** (2010), no. 2, 347–362.
- [8] I. T. Degirmenci, *Asymptotic analysis and performance-based design of large scale service and inventory systems*, Ph.D. dissertation, Department of Business Administration, Duke University (2010).

- [9] J. M. Dobbie, *A doubled-ended queuing problem of Kend*, Operations Research **9** (1961), no. 5, 755–757.
- [10] Stewart N. Ethier and Thomas G. Kurtz, *Markov processes : characterization and convergence*, J. Wiley & Sons, 1986.
- [11] O. Garnet, A. Mandelbaum, and M. Reiman, *Designing a call center with impatient customers*, Manufacturing & Service Operations Management **4** (2002), no. 3, 208–227.
- [12] S. M. Giveen, *A taxicab problem with time-dependent arrival rates*, SIAM Reviw **5** (1963), no. 2, 119–127.
- [13] Donald L. Iglehart and Ward Whitt, *The equivalence of functional central limit theorems for counting processes and associated partial sums*, The Annals of Mathematical Statistics **42** (1971), no. 4, 1372–1378.
- [14] J. Jacod and A.N. Shiryaev, *Limit theorem for stochastic processes*, 2 ed., Springer-Verlag, Berlin, 2003.
- [15] Weining Kang and Kavita Ramanan, *Fluid limits of many-server queues with reneging*, Ann. Appl. Prob. **20** (2010), no. 6, 2204–2260.
- [16] B.R.K. Kashyap, *The double-ended queue with bulk service and limited waiting space*, Operations Research **14** (1966), no. 5, 822–834.
- [17] D. G. Kendall, *Some problems in the theory of queues*, Journal of the Royal Statistical Society. Series B **13** (1951), no. 2, 151–185.
- [18] W. K. Kim, K. P. Yoon, G. Mendoza, and M. Sedaghat, *Simulation model for extended double-ended queueing*, Computers & Industrial Engineering **59** (2010), no. 2, 209–219.
- [19] C. Lee and A. Weerasinghe, *Convergence of a queueing system in heavy traffic with general patience-time distributions*, Stochastic Processes and their Applications **121** (2011), no. 11, 2507–2552.
- [20] Xin Liu, Qi Gong, and Vidyadhar G. Kulkarni, *Diffusion models for double-ended queues with renewal arrival processes*, Stoch. Syst. **5** (2015), no. 1, 1–61.
- [21] A. Mandelbaum and P. Momcilovic, *Queues with many servers and impatient customers*, Mathematics of Operations Research **37** (2012), no. 1, 41–65.
- [22] Guodong Pang, Rishi Talreja, and Ward Whitt, *Martingale proofs of many-server heavy-traffic limits for markovian queues*, Probab. Surveys **4** (2007), 193–267.
- [23] D. Perry and W. Stadje, *Perishable inventory systems with impatient demands*, Mathematical methods of operations research **50** (1999), no. 1, 77–90.
- [24] B. Prabhakar, N. Bambos, and T. S. Mountford, *The synchronization of Poisson processes and queueing networks with service and synchronization nodes*, Advances in Applied Probability **32** (2000), no. 3, 824–843.
- [25] J. Reed and T. Tezcan, *Hazard rate scaling of the abandonment distribution for the  $GI/M/n + GI$  queue in heavy traffic*, Oper. Res. **60** (2012), no. 4, 981–995.
- [26] J. E. Reed and Amy R. Ward, *Approximating the  $gi/gi/1+gi$  queue with a nonlinear drift diffusion: Hazard rate scaling in heavy traffic*, Math. Oper. Res. **33** (2008), no. 3, 606–644.
- [27] A. R. Ward and P. W. Glynn, *A diffusion approximation for a Markovian queue with reneging*, Queueing Syst. Theory Appl. **43** (2003), no. 1/2, 103–128.
- [28] ———, *A diffusion approximation for a  $GI/GI/1$  queue with balking or reneging*, Queueing Syst. Theory Appl. **50** (2005), no. 4, 371–400.
- [29] S. Zeltyn and A. Mandelbaum, *Call centers with impatient customers: Many-server asymptotics of the  $M/M/n + G$  queue*, Queueing Systems **51** (2005), no. 3-4, 361–402.
- [30] S. A. Zenios, *Modeling the transplant waiting list: A queueing model with reneging*, Queue-



XIN LIU  
DEPARTMENT OF MATHEMATICAL SCIENCES  
CLEMSON UNIVERSITY  
CLEMSON, SC 29634, USA  
EMAIL: XLIU9@CLEMSON.EDU.